The Evaluation of Industrial Application of Fourier Transform Infrared (FT-IR) Spectroscopy and Multivariate Data Analysis Techniques for Quality Control and Classification of South African Spirit Products

by
Tania Victoria Kleintjes

Thesis presented in partial fulfilment of the requirements for the degree of
Master of Agricultural Science

at
Stellenbosch University
Department of Viticulture and Oenology, Faculty of AgriSciences

Supervisor: Professor Marius Lambrechts
Co-supervisor: Doctor Hélène Nieuwoudt

December 2013

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 2 October 2013

Summary

The WineScan FT120 is widely used in wine laboratories across South Africa. The WineScan FT120 uses Fourier transform infrared (FT-IR) spectroscopy with multivariate data analysis to correlate spectra with chemical compositional data. Ready-to-use, commercially available calibration models for a FT-IR spectroscopy instrument are an advantage for unskilled users and routine analysis. Introducing spirit products to this technology introduced new interferences, which necessitated vastly different calibrations models to compensate for the changes.

Accuracy, precision and ruggedness of the reference methods validated during method validation, verified the suitability of the reference methods used to quantify the parameters in question before calibration model building was attempted.

Various principal component analysis (PCA) were performed prior to the calibration step with the aim to identify outliers and inspect groupings. PCA models could identify samples with atypical spectra and differentiate between product types.

Two tactics regarding data sets for calibration set-up was experimented with, all the products together and calibration models per product. Partial least squares (PLS) regression was used to establish the calibration models for ethanol, density, obscuration and colour. With all the calibration models, the calibration models based on the product specific data sets, achieved better predicting statistics. The best performing ethanol calibration models achieved Residual mean square error of prediction (RMSEP) = 0.038 to 0.106 %v/v and showed significant improvement on previously reported prediction errors by Lachenmeier (2007). The results for the density calibration showed a similar trend, with the product specific calibration models outperforming the calibration model when all samples were included into one calibration model. This study produced novel results for quantification of obscuration (RMSEP = 0.10 and 0.09 in blended brandies and potstill brandies, respectively) and colour (RMSEP < 2.286 gold units) of brandies and whiskies. The correlation coefficients (R^2) between true and predicted values, for the four parameters tested, indicated good to excellent precision ($0.8 < R^2 < 1.0$). Minimising the variation between the samples of the data set, gave more accurate regression statistics, but this resulted in a lower residual predictive deviation (RPD) value (< 5) that indicated models were not suitable for quantification. Adding more samples per product will add more variability into a data set per product, increase the SD and result in an increase in the RPD. The results pave the way for the development of calibration models for the quantification of other parameters for specific products.

Following the groupings of product types, further classifications of brandy brands were investigated. PCA plots showed clear separation between potstill brandies and blended brandies and some degree of clustering between some of the blended brands was observed. Classification of brandies were investigated using the Soft Independent Modeling of Class Analogy (SIMCA) approach resulting in a total correct classification rates between 81.25% and 100% for the various brandy brands. These preliminary results were very promising and highlight the potential of using FT-IR spectroscopy and multivariate classification techniques as a tool for rapid quality control and authentication of brandy brands.

Using this work as base for further classification projects, this could be of great benefit to the alcoholic beverage industry of South Africa. Future work will involve the development of a database comprised of more products guaranteed authentic to expand the discriminating options. The results suggest FT-IR spectroscopy could be useful in authentication studies.

Opsomming

Die WineScan FT120 is ‘n algemeen gebruikte instrument regoor Suid-Afrika. Die WineScan FT120 gebruik Fourier-transformasie-infrarooi (FT-IR) spektroskopie tesame met multiveranderlike statistiese metodes om spektra te korreleer met chemiese samestellingsdata. Die kommersieël beskikbare kalibrasiemodelle vir die FT-IR spektroskopie-instrument is ‘n voordeel vir onbedrewe gebruikers en roetine ontleding. Blootstelling van spiritusprodukte aan die tegnologie, het nuwe hindernisse bekend gestel en dus is verskillende kalibrasiemodelle genoodsaak om hiervoor te kompenseer.

Akkuraatheid, presiesheid en ruheid van die verwysingsmetodes is geëvalueer tydens metodevalidasie. Die verwysingsmetodes is geskik verklaar vir die konstruksie van die kalibrasiemodel met geverifieërde akkurate verwysingsresultate.

Verskeie multiveranderlike hoofkomponentanalise (MVK) was uitgevoer voor die kalibrasie-stap met die doel om uitskieters te identifiseer en groeperings te inspekteer. MVK modelle kon monsters met atipiese spektra identifiseer en onderskei tussen verskillende produk tipes.

Twee taktieke aangaande datastelsamestelling is getoets tydens kalibrasiemodel-opstelling, al die produkte saam en kalibrasiemodelle per produk soos met die MVK aangedui. Parsiële kleinste kwadraat (PKK)- regressie is gebruik vir die opstel van die kalibrasiemodelle vir etanol, digtheid, obskurasie en kleur. Met al die kalibrasiemodelle het die produk spesifieke kalibrasiemodelle beter regressiestatistiek gelewer. Die beste presterende etanol kalibrasiemodelle het ‘n standaardvoorspellingsfout (SVF) = 0.038 tot 0.106 %v/v bereik en het ‘n beduidende verbetering getoon op vorige gerapporteerde studies op spiritusprodukte (Lachenmeier, 2007). Die resultate vir die digtheidskalibrasiemodelle het ‘n eenderse tendens getoon soos die etanol, met die produk spesifieke kalibrasiemodelle wat beter presteer het. Hierdie studie was eerste in sy soort met die kalibrasiemodel vir obskurasie (SVF = 0.10 en 0.09 in gemengde brandewyne en potketel brandewyne, onderskeidelik) en kleur (SVF < 2.286 goud eenhede) van brandewyne en whiskies. Die bepalingskoëffisiënt (R^2) vir die vier parameters, dui op goeie tot uitstekende presiesheid ($0.8 < R^2 < 1.0$). Vermindering van die variasie tussen die monsters in die datastel, het meer akkurate regressiestatistiek teweeg gebring, maar ‘n laer relatiewe voorspellingsafwyking (RVA) waarde (<5) tot gevolg gehad wat aan dui dat hierdie modelle nie geskik is vir sifting of kwantifisering nie. Die byvoeging van meer monsters per produk sal meer verskeidenheid in die datastel per produk bring, wat dan die standaardafwyking sal laat toeneem en uiteindelik die RVA laat toeneem. Die resultate het die fondasie gelê vir die ontwikkeling van kalibrasiemodelle vir die kwantifisering van ander parameters vir spesifieke produkte.

As opvolg tot die groeperings van die produk tipe, waargeneem in die MVK modelle, was klassifikasie van brandewyn handelsmerke ondersoek. MVK modelle het duidelike skeiding gewys tussen potketel en gemengde brandewyne en tot ‘n sekere mate groepering tussen handelsmerke. Klassifikasie van brandewyne was ondersoek met behulp van the Soft Independent Modeling of Class Analogy (SIMCA) met die resultaat van ‘n totale korrekte klassifikasiekoers van tussen 81.25% en 100% vir die verskeie brandewyn handelsmerke. Hierdie voorlopige resultate toon belowend en beklemtoon die potensiaal van FT-IR spektroskopie en chemometrics tegnieke as toerusting vir die vinnige kwaliteitskontrole en egtheid van brandewyn handelsmerke studies.

Met hierdie werk as basis vir verdere klassifikasie projekte, kan dit ‘n groot aanwins wees tot die alkoholiese drank industrie van Suid-Afrika. Toekomstige werk sal insluit die ontwikkeling van ‘n databasis saamgestel met meer gewaarborgde egte produkte om die klassifikasie uit te brei.

This thesis is dedicated in loving memory of my mother for being the driving force behind everything I do and to my family and husband for their love and support.

Acknowledgements

I wish to express my sincere gratitude and appreciation to the following persons and institutions:

- **PROFESSOR MARIUS LAMBRECHTS,** my project leader at Distell, South Africa for believing in me when everyone else has given up. Thank you for the support and guidance throughout this study.

- **DR HÉLÈNE NIEUWOUDT**, Institute for Wine Biotechnology, Stellenbosch University, my co-supervisor, for her guidance, suggestions, support and criticism when it was most needed.

- **MY HUSBAND, SHAMUS,** who gave me his unwavering support and encouragement and for putting up with months of neglect.

- **MY FAMILY, FRIENDS AND COLLEAGUES,** for all the support, love, encouragement and patience throughout this study.

- **DISTELL, SOUTH AFRICA** for supplying the samples and unrestricted use of laboratory facilities, especially the use of the WineScan FT120 at the Adam Tas site. Especially the laboratory staff at Adam Tas for their help in the wet analysis of this study and patience with this student in their busy laboratory. A special mention to **ELANA LOCHNER** for guidance and help during the sampling (which spanned over years) of this study.

- **DR CATRINE DE VILLIERS**, for her steadfast support and kick-start attitude throughout this process.

- **THE ALMIGHTY,** for my gifts and blessings.

Preface

This thesis is presented as a compilation of six chapters. Each chapter is introduced separately.

Chapter 1 **General introduction and project aims**

Chapter 2 **Literature review**
INTRODUCTION TO SOUTH AFRICAN SPIRIT PRODUCTS, CURRENT METHODS OF ANALYSES AND DATA ANALYSIS; REVIEW OF THE APPLICATIONS OF INFRARED SPECTROSCOPY IN THE ALCOHOLIC BEVERAGE INDUSTRY

Chapter 3 **Research results**
VALIDATION OF METHOD OF ANALYSIS FOR ETHANOL, DENSITY, OBSCURATION AND COLOUR OF SPIRIT PRODUCTS IN A COMMERCIAL LABORATORY IN SOUTH AFRICA

Chapter 4 **Research results**
QUANTIFICATION OF ETHANOL, DENSITY, OBSCURATION AND COLOUR OF SPIRIT PRODUCTS USING FOURIER TRANSFORM INFRARED (FT-IR) SPECTROSCOPY AND MULTIVARIATE DATA ANALYSIS METHODS

Chapter 5 **Research results**
FEASIBILITY STUDY OF CLASSIFICATION OF BRANDY PRODUCTS USING FOURIER TRANSFORM INFRARED (FT-IR) SPECTROSCOPY AND MULTIVARIATE DATA ANALYSIS

Chapter 6 **General discussion and conclusions**

Contents

CHAPTER 1. GENERAL INTRODUCTION AND PROJECT AIMS 1

1.1 Introduction 2
1.2 Project aims 4
1.3 References 5

CHAPTER 2. LITERATURE REVIEW: INTRODUCTION TO SOUTH AFRICAN SPIRIT PRODUCTS, CURRENT METHODS OF ANALYSES AND DATA ANALYSIS; REVIEW OF THE APPLICATIONS OF INFRARED SPECTROSCOPY IN THE ALCOHOLIC BEVERAGE INDUSTRY 7

2.1 Introduction 8
2.2 Production of spirits products 8
2.2.1 Distillation 9
2.2.2 Brandy 9
2.2.3 Whisky 10
2.2.4 Vodka 10
2.2.5 Gin 10
2.3 Parameters of importance to this study 11
2.3.1 Ethanol 11
2.3.2 Density 12
2.3.3 Obscuration 12
2.3.4 Colour 13
2.4 Analytical method validation 13
2.5 Univariate and multivariate data analysis 15
2.5.1 Univariate data analysis 15
2.5.2 Multivariate data analysis 15
2.5.2.1 Data exploration/description 16
2.5.2.2 Regression and prediction 16
2.5.2.3 Discrimination and classification 17
2.5.2.4 Outlier detection 17
2.6 Infrared Spectroscopy 18
2.6.1 A review of quantitative studies with IR spectroscopy in the alcoholic beverage industry 18
2.6.2 A review of qualitative studies with IR spectroscopy in the alcoholic beverage industry 25
2.7 Conclusion 28
2.8 References 29

CHAPTER 3. VALIDATION OF METHOD OF ANALYSIS FOR ETHANOL, DENSITY, OBSCURATION AND COLOUR OF SPIRIT PRODUCTS IN A COMMERCIAL LABORATORY IN SOUTH AFRICA 34

Abstract 35
3.1 Introduction 36
3.2 Materials and methods 37
3.2.1 Sample collection 37
3.2.2 Methods validated 37
3.2.2.1 Ethanol concentration 37
3.2.2.2 Density 38
3.2.2.3 Obscuration 38
3.2.2.4 Colour 38
3.2.3 Statistical analysis 39
3.3 Results and discussion 39
3.3.1 Ethanol analysis method validation 39
3.3.2 Density analysis method validation 41
3.3.3 Obscuration analysis method validation 42
3.3.4 Colour analysis method validation 43
3.4 Conclusion 44
3.5 References 45

CHAPTER 4. QUANTIFICATION OF ETHANOL, DENSITY, OBSCURATION AND COLOUR OF SPIRIT PRODUCTS USING FOURIER TRANSFORM INFRARED (FT-IR) SPECTROSCOPY AND MULTIVARIATE DATA ANALYSIS METHODS 47

Abstract 48
4.1 Introduction 49
4.2 Materials and methods 50
4.2.1 Samples 50
4.2.2 Reference methods 50
4.2.3 FT-IR spectral acquisition 50
4.2.4 Data analysis 51
4.2.4.1 Explorative data analysis 51
4.2.4.2 PLS regression modelling 51
4.2.4.3 Statistics used to evaluate the calibration models 52
4.3 Results and discussion 52
4.3.1 Overview of the data set 52
4.3.2 Evaluation of the spectra 53
4.3.3 Explorative data analysis 54
4.3.4 PLS regression modelling 58
4.3.4.1 Ethanol PLS regression models 62
4.3.4.2 Density PLS regression models 63
4.3.4.3 Obscuration PLS regression models 64
4.3.4.4 Colour PLS regression models 65
4.4 Conclusion 66
4.5 References 67

CHAPTER 5: FEASIBILITY STUDY OF CLASSIFICATION OF BRANDY PRODUCTS USING FOURIER TRANSFORM INFRARED (FT-IR) SPECTROSCOPY AND MULTIVARIATE DATA ANALYSIS 70

Abstract 71
5.1 Introduction 72
5.2 Materials and methods 73
5.2.1 Samples 73
5.2.2 Spectra 73
5.2.3 Multivariate data analysis 73
5.3 Results and discussion 74
5.3.1 Explorative data analysis 74
5.3.2 Classification of brandy samples 77
5.4 Conclusion 82
5.5 References 83

CHAPTER 6: GENERAL DISCUSSION AND CONCLUSIONS 85

General discussion and conclusions 86
References 88

Chapter 1

General Introduction and project aims

CHAPTER 1: General Introduction and project aims

1.1 INTRODUCTION

Brand is a valuable asset to companies, with no exception to the alcoholic beverage industry since brands convey competence, quality and image to the consumer (O'Cass *et al.,* 2000). Consumers expect a product of consistent quality, in other words, a product with visual and sensorial consistency. South Africa is an internationally recognised producer of high quality, prize winning (IWSC, 2012) spirit products. Brandy is South Africa's biggest exported spirit product with a total of 791 697 L exported in 2012 (SAWIS, 2012), while whisky is the biggest imported spirit product with 15 330 969 L imported in 2012 (SAWIS, 2012). To remain an international competitive producer in spirit products, and also increase its reputation as a producer of quality spirit products, South Africa needs to be up-to-date with the new technologies and developments within the science arena. Internal product and process specifications are established throughout production processes, to monitor the production and ensure that requirements are met. Quality control to ensure compliance to specifications plays a fundamental role in protecting the integrity of a brand, while providing brand authentication and protection against counterfeit products to keep the consumers' loyalty.

The spirit product parameters investigated in this study were ethanol, density, obscuration and colour. The product types used in this study were brandy, whisky, vodka and gin. The South African Liquor Products Act (Act 60 of 1989) stipulates the legal ethanol concentration for each spirit product type. Taxes and penalties that are imposed ensure compliance from producers. Density plays an integral role in the determination of accurate volume. Producers are audited on legal requirements, stipulated in SANS 1841 (2008) in South Africa, on the average volume during bottling. The total extract of spirits, in this study only done for brandy, is normally expressed as the obscuration. Obscuration is of organoleptic importance as it is a measure of sweetness, adding to the mouth-feel of a product. Colour measurements are part of the final quality check of a product. Brandy and whisky products were tested for this parameter. Consistency in colour, specific to each product must be maintained. Accurate, precise and repeatable measurements of these parameters are thus of utmost importance to spirit producers. Current methods of analysis of these parameters do not fulfil the needs of a fast paced production environment. At the same time, protecting brands against unscrupulous producers in authentication studies is needed. Infrared spectroscopy coupled with multivariate data analysis, offer such a solution. This methodology can quantify various parameters within a single sample, in the fraction of the time required by the traditional reference techniques, and provides a database of extremely valuable information per sample that can be used as chemical fingerprints to confirm the identity of a sample (Cozzolino, 2012).

The use of infrared spectroscopy is well documented in South African wine laboratories as a method of analysis of glycerol in wine (Nieuwoudt *et al.,* 2004), a screening method for fermentation products of *Saccharomyces* yeasts in Chenin blanc fermentations (Nieuwoudt *et al.,* 2006), determination of total phenolics and total anthocyanins of grapes (Lochner, 2006), amongst others. The quantitative application of infrared spectroscopy to spirit products is well documented. It has been used in ethanol quantification of beer to spirit products (Gallignani *et al.,* 1994), non-invasive determination of ethanol in whiskies and vodka samples (Nordon *et al.,* 2005) and in a multi-component quantification study on spirits and beer samples (Lachenmeier, 2007). Despite these successful applications of infrared spectroscopy on spirit products, there have not been quantitative studies done on South African spirit products.

The use of infrared spectroscopy in classification studies were documented in studies by Picque *et al.* (2006) on discrimination of cognacs, Pontes *et al.* (2006) on the verification of adulteration of spirit products by water, ethanol and methanol, and in the detection of counterfeit whiskies by McIntyre *et al.* (2011). South African brandies have previously been included in a classification study where it was shown that brandies could be differentiated on the basis of country of origin (Palma & Barroso, 2002). However, to date there has been no reported study on the brand validation of South African spirit products, using infrared spectroscopy and multivariate techniques. Both producers and consumers can benefit from the classification of SA spirit products that will aid traceability and ensure a consistent and conforming product, year after year.

Modern analytical instruments have given chemists access to masses of information. The need to digest this information into sensible results has driven the development of multivariate data analysis techniques that use mathematical and statistical procedures to extract the maximum useful information from the data (Berrueta *et al.,* 2007). Partial Least Squares (PLS) regression is a method to relate the variations in spectra to the variations seen in the chemical parameters (Esbensen, 2002). PLS regression has been used for developing calibration models for the determination of quantitative parameters in various spirit products studies, such as the determination of relative density, ethanol concentration and total dry extract in spirits and liqueurs (Arzberger and Lachemeier, 2008). Principal Component Analysis (PCA) is an often used modelling method used for interpretation of spectral data. This method allows for the visualisation of the relationship between variables, determination and interpretation of sample patterns, groupings, similarities and/or differences (Esbensen, 2002). Previously mentioned classification studies (Pontes *et al.,* 2006; Palma & Barroso, 2002; Picque *et al.,* 2006; McIntyre *et al.,* 2011) have used PCA. Soft Independent Modeling of Class Analogy (SIMCA) is an often used class-modeling technique (Berrueta *et al.,* 2007) and is based on individual PCA models created for each class of samples. SIMCA has been used in the classification and verification of adulteration in whisky, brandy, rum and vodka study with very good success rate (Pontes *et al.,* 2006).

WineScan FT120 is a multicomponent analytical instrument used in various studies on wine in South Africa (Nieuwoudt *et al.,* 2004; Lochner, 2006; Magerman, 2006), but it has not been evaluated for use on spirit products. The WineScan FT120 was the first purpose-built wine analyser (Foss Analytical, Denmark) to use the mid-infrared region (400 to 5000 cm^{-1}), since this region offers more accurate determination of compounds than near-infrared spectroscopy (Patz *et al.,* 2004) and acquires spectra from 926 to 5012 cm^{-1} for each sample. This region includes the fingerprint region, 983 – 1149 cm^{-1}, which contain a significant amount of variation related to the absorbance by molecular chemical groups present in alcoholic beverages (Nieuwoudt *et al.,* 2004; Alonso-Simón *et al.,* 2004).

Despite the obvious advantages of this technique, namely speed, simplicity and potential for classification, no studies were reported on the application thereof on South African spirit products. In this study, the WineScan FT120 instrument was used to generate the spectra of various types and brands of spirit products. The spectra were correlated with PLS regression methods to quantify ethanol, density, obscuration and colour of spirit products. PCA was used to identify the similarities and/or differences between the products and types, while the SIMCA approach was used to classify nine individual brandy brands. These classification results were used as a preliminary result and first step towards authentication of brandy brands.

1.2 PROJECT AIMS

The project aims were formulated based on the shortcomings identified in the introductory section and established foundation of the mentioned methodologies. The aims of this project were therefore twofold: (i) to quantify ethanol, density, obscuration and colour in SA brandy, whisky, gin and vodka in a routine commercial laboratory by FT-IR spectroscopy, and (ii) to classify brandy products according to their brands. Aspects that were investigated were:

- Evaluation of the precision and verification of the accuracy of the reference analysis methods used in the routine laboratories for these parameters (Chapter 3);
- Development of PLS calibration models for the WineScan FT120 instrument for the quantification of ethanol, density, obscuration and colour in spirit products (Chapter 4).
- Development of classification models for nine brandy brands, as preliminary steps towards authentication of brandy brands (Chapter 5).

1.3 REFERENCES

1. Arzberger, U. & Lachenmeier, D.W. (2008). Fourier Transform Infrared Spectroscopy with Multivariate analysis as a Novel Method for Characterizing Alcoholic Strength, Density and Total Dry Extract in Spirits and Liquers. *Food Analytical Methods* **1**, 18 – 22.
2. Alonso-Simón, A., Encina, A.E., García-Angulo, P., Álvarez, J.M. & Acebes, J.L. (2004). FTIR spectroscopy monitoring of cell wall modifications during the habituation of bean (*Phaseolus vulgaris* L.) callus cultures to dichlobenil. *Plant Science,* **167**, 1273 – 1281.
3. Berrueta, L.A., Alonso-Salces, R.M. & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A* **1158**, 196-214.
4. Cozzolino, D. (2012). Recent Trends on the Use of Infrared Spectrosocpy to Trace and Authenticate Natural and Agricultural Food Products. *Applied Spectroscopy Reviews,* **47**, 518 – 530.
5. Esbensen, K.H., (2002). Multivariate Data Analysis – In Practice. An introduction to multivariate data analysis and experimental design, *5th ed*. Camo ASA, Oslo.
6. Gallignani, M., Garrigues, S. & De la Guardia, M. (1994). Derivative Fourier transform infrared spectrometric determination of ethanol in alcoholic beverages. *Analytica Chimica Acta* **287,** 275-283.
7. Lachenmeier, D.W. (2007). Rapid quality control of spirit drinks and beer using multivariate data analysis of Fourier transform infrared spectra. *Food Chemistry* **101,** 825-832.
8. Lochner, E. (2006). The evaluation of Fourier transform infrared spectroscopy (FT-IR) for the determination of total phenolics and total anthocyanins concentration of grapes, M.Sc. (Agriculture) thesis, Stellenbosch University.
9. Magerman, C.M. (2009). The evaluation of Fourier transform infrared (FT-IR) spectroscopy for quantitative and qualitative monitoring of alcoholic wine fermentation, M.Sc. (Wine Biotechnology) thesis, Stellenbosch University.
10. McIntyre, A.C., Bilyk, M.L., Nordon, A., Colquhoun, G. & Littlejohn, D. (2011). Detection of counterfeit Scotch whisky samples using mid-infrared spectrometry with an attenuated total reflectance probe incorporating polycrystalline silver halide fibres. *Analytica Chimica Acta* **690**, 228–233.
11. Nieuwoudt, H.H., Prior, B.A., Pretorius, I.S., Manley, M. & Bauer, F.F. (2004). Principal Component Analysis Applied to Fourier Transform Infrared Spectroscopy for the Design of Calibration Sets for Glycerol Prediction Models in Wine and for the Detection and Classification of Outlier Samples. *Journal of Agricultural and Food Chemistry* **52**, 3726 – 3735.
12. Nieuwoudt, H.H., Pretorius, I.S., Bauer, F.F., Nel, D.G. & Prior, B.A. (2006). Rapid screening of the fermentation profiles of wine yeasts by Fourier transform infrared spectroscopy. *Journal of Microbiological Methods* **67,** 248-256.
13. Nordon, A., Mills, A., Burn, R.T., Cusick, F.M. & Littlejohn, D. (2005). Comparison of non-invasive NIR and Raman spectrometries for determination of alcohol content of spirits. *Analytica Chimica Acta* **548,** 148-158.
14. O'Cass, A., Lim, K. & Julian, C.C. 2000. Brand classifications: identifying the origin of brands. *Visionary marketing for the 21st Century: facing the challenge: Proceedings of the Australian and New Zealand Marketing Academy (ANZMAC) Conference,* Gold Coast, Queensland, 28 November – 1 December, Griffith University Press, 871 – 878.

15. Palma, M. & Barroso, C.G. (2002). Application of FT-IR spectroscopy to the characterisation and classification of wines, brandies and other distilled drinks. *Talanta* **58,** 265-271

16. Picque, D., Lieben, P., Corrieu, G., Cantagrel, R., Lablanquie, O. & Snakkers, G. (2006). Discrimination of Cognacs and other distilled drinks by Mid-Infrared Spectroscopy. *Journal of Agricultural and Food Chemistry* **54**, 5220-5226.

17. Pontes, M.J.C, Santos, S.R.B., Araújo, M.C.U., Almeida, L.F., Lima, R.A.C., Gaião, E.N. & Souto, U.T.C.P. (2006). Classification of distilled alcoholic beverages and verification of adulteration by near infrared spectrometry. *Food Research International* **39,** 182-189.

18. South African Liquor Products Act No. 60 of 1989.

19. South African National Standard. SANS 1841. (2008). Control of the quantity of contents in prepacked packages within the prescriptions of legal metrology legislation.

20. IWSC. (2012). International Wine and Spirits Competition. [URL. http://www.iwsc.net/home]. Cited the 09th of December 2012.

21. SAWIS. (2012). South African wine industry information and systems. [URL. http://www.sawis.co.za/info/statistics.php]. Cited the 09th of December 2012.

Chapter 2

Literature review:

Introduction to South African spirit products, current methods of analyses and data analysis; Review of the applications of infrared spectroscopy in the alcoholic beverage industry

CHAPTER 2: Introduction to South African spirit products, current methods of analyses and data analysis; Review of the applications of infrared spectroscopy in the alcoholic beverage industry

2.1 INTRODUCTION

The South African spirits industry is well established and internationally recognised for the quality of its products. This is illustrated by the fact that South Africa won the International Wine and Spirits Competition (IWSC) title for the best brandy in the world, nine times in the past 13 years (IWSC, 2012). From a socio-economic perspective, the activities involved in the production and distribution of spirits generate employment and provide an important source of tax revenue for the government (Punt, 2010). Statistics show that the spirits and ready-to-drink (RTD) industries support an excess of 54 000 jobs throughout the South African economy (Punt, 2010). Total sales volumes of alcoholic beverages in South Africa for the 2008/2009 financial year, amounted to 3.3 billion litres, amounting to an estimated R57.5 billion. The main role players in terms of volume share of spirits sold in South Africa, include Distell Group (Pty) Ltd (40.5%), Brandhouse Beverages (Pty) Ltd (22%), and Edward Snell & Co (16%). Other role players in the spirits and RTD industries are South African Breweries (SAB) millers, DGB, KWV, Pernod Ricard, NMK Schulz, NCP Alcohols and Illovo (Punt, 2010).

In this fast moving industry producers require accurate, precise, repeatable and available on demand measurements. Products need to conform to internal specifications, measure up to the requirements of the consumers and obey legal specifications (South African Liquor Products Act 60 of 1989). Throughout the production process various specifications must be adhered to, to ensure that a uniform product is produced year after year from seasonal raw materials.

Current methods of chemical analysis in spirit products are laborious, time consuming and susceptible to error. Infrared spectroscopy, coupled with multivariate data analysis techniques, a well-established and proven method of analysis in the wine industry, was evaluated for the purpose of this review, for their application in the spirits. With this summary of the studies done in the alcoholic beverage industry, the interest, success and shortcomings of the technique in in this industry are demonstrated.

2.2. PRODUCTION OF SPIRITS PRODUCTS

The production of spirit products is closely regulated by The Liquor Products Act (South African Liquor Products Act 60 of 1989). Spirits are produced by the distillation of a fermented base product. The raw material and different classes of spirits, according to the South African Liquor Products Act 60 of 1989, are given in Table 2.1.

Table 2.1: Raw material and different classes of spirits according The Liquor Products Act 60 of 1989 [Subreg. (1) amended by GN R21/2001, GN R77/2006 and GN R555/2009]

Raw material	Class of spirit
The fermented product of the vine	grape spirit; husk spirit; premium husk spirit; pot still brandy; blended brandy; vintage brandy
The fermented mash of grain or malt	whisky; malt whisky; blended whisky
Fermented sugar cane juice, sugar cane syrup or sugar cane molasses	cane spirit; rum
Any fermented harmless vegetable article	gin; vodka; unspecified spirit; mixed spirit

2.2.1 Distillation

The distillation process concentrates ethanol and eliminates some of the by-products produced during fermentation. Two types of distillation methods are used in spirits production: potstill and continuous, respectively. Potstill distillation is a 'batch process' whereby both the ethanol and flavours present in the base wine are concentrated (Le Roux, 1997).

With potstill distillation, wine used for distilling with an ethanol content of 10 - 12 %v/v is distilled into low wine, with an ethanol content of approximately 30 %v/v. The first distillation may be regarded as a concentration process and lasts about six to eight hours. The volatile compounds are separated from the non-volatile components and most of the water present. The second stage is the distillation of the low wine into three fractions in sequence, namely the heads, the hearts and the tails, with ethanol content usually between 68 and 70 %v/v. The second distillation is a slower process and lasts about 12 to 14 hours. The heads are usually collected within the first 15 minutes of the process. The heart fraction contains all the positive flavour compounds that contribute to the aroma and flavour of brandy and whisky. The tails fraction contains quantities of high-boiling fatty acids that would impart soapy, fatty and oily characteristics to spirits if left to distill in the heart fraction (Le Roux, 1997). Spirits with distinctive flavours such as brandy and whisky are produced with potstill distillation.

Continuous distillation is performed in column stills that concentrate only the ethanol, while removing all other flavours present in the fermented raw material and results in neutral spirits. In a continuous distillation process, fermented product is continually fed into the first column while the neutral spirit is continually and simultaneously drawn from another column. Fractions of volatile compounds are drawn at several points in the system. These fractions are made possible by the differing boiling points of the flavour compounds present. The spirit drawn from the final column has an ethanol concentration of 96.4 %v/v and has no odour or flavours (Le Roux, 1997). This system is favoured for the production of vodka and gin.

Requirements for the different products are given below.

2.2.2 Brandy

According the requirements of The SA Liquor Products Act (Act 60 of 1989), a potstill brandy must be distilled from the fermented juice of the product of the vine, which is distilled under excise supervision in a pot still to an ethanol content of not more than 75 %v/v. It must be matured for at least three years in oak barrels, with a capacity of not more than 340 L and that have been approved by the Commissioner of Customs and Excise or a person authorised thereto by him or her in writing. The potstill brandy must be approved by the board and be certified by it as a spirit produced exclusively from the fermented juice of the product of the vine and must have an ethanol content of at least 38 %v/v (South African Liquor Products Act 60 of

1989). Potstill brandy is the richest and most layered in flavour of the three types of brandy produced in South Africa (Le Roux, 1997). The final product must contain a minimum of 90% potstill brandy matured for at least three years, and a maximum of 10% neutral wine spirit from column-still distillation (unmatured) (South African Liquor Products Act 60 of 1989). Commercial examples used in this study include Klipdrift Gold and Flight of the Fish Eagle.

The requirements for blended brandy according The SA Liquor Products Act 60 of 1989 are that blended brandy must consist of a mixture of not less than 30% potstill brandy (matured for at least three years) to which no grape spirit, wine spirit, spirit or a mixture thereof has been added and not more than 70% neutral wine spirit from the column-still distillation (unmatured). It must have an ethanol content of at least 43 %v/v (South African Liquor Products Act 60 of 1989). Commercial examples used in this study include Viceroy, Klipdrift Export and Richelieu.

According The Liquor Products Act 60 of 1989 a vintage brandy shall be produced in such a manner that at least 90% of the volume thereof is blended brandy of which the portion that must be aged, matures for a further five years in oak barrels with a capacity of not more than 1 000 liters; and the other portion has also been matured for at least eight years in oak barrels and must have an ethanol content of at least 38 %v/v (South African Liquor Products Act 60 of 1989). Since vintage brandies have to be aged for a minimum of eight years and contain a significant portion of matured wine spirit, vintage brandy has a distinctive wood maturation character when compared to potstill and blended brandy (Le Roux, 1997).

2.2.3 Whisky

The distillation of fermented mash of grain or malt yields whisky. According The Liquor Products Act (Act 60 of 1989) whisky must be produced from a mash of grain in which the diastase of the malt contained therein, has brought about sugar conversion, which has been fermented by the activity of yeast and must have been distilled at less than 94,8 %v/v ethanol content. The distilled liquid must be matured for at least three years in wooden casks approved for this purpose by the Commissioner of Customs and Excise with a capacity of not more than 700 L and have an ethanol content of at least 43 %v/v (South African Liquor Products Act 60 of 1989). The commercial examples used in this study are Harrier Whisky and Three Ships Whisky.

2.2.4 Vodka

Vodka must be produced by the distillation of any fermented, harmless vegetable article, must not have any distinctive characteristic, aroma, taste or colour; and have an ethanol content of at least 43 %v/v (South African Liquor Products Act 60 of 1989). The commercial example used in this study is Romanoff.

2.2.5 Gin

Gin shall be produced by the distillation of the fermented mash of grain together with juniper berries and must have an ethanol content of at least 43 %v/v (South African Liquor Products Act 60 of 1989). The commercial examples of gin used in this study are Old Buck Gin and Gordon's Gin.

2.3 PARAMETERS OF IMPORTANCE TO THIS STUDY

2.3.1 Ethanol

The ethanol content of spirit products is a key analytical parameter in the distilled beverage industry. It plays an integral role from the onset of the process of distilled products production. Ethanol also determines the quality and preservation of the product (González-Rodríguez *et al*., 2003). The economic implications with regards to taxes and regulations by governing bodies make ethanol an important parameter for producers (South African Liquor Products Act No. 60 of 1989; SANS 1841, 2008).

The methods used in South African laboratories to determine ethanol concentration in alcoholic beverages are distillation, infrared spectroscopy, boiling point and enzymatic methods.

The distillation methods for ethanol measurement, determine the density of the spirit after distillation and subsequently the ethanol concentration by pycnometry that involves the accurate determination of mass and volume of the distillate to obtain its density and subsequently the ethanol (Brereton *et al*., 2003). The density of a sample (distillate) can be determined by the following equation:

$$\text{Density of distillate} = E - A / V$$

where: *A* denotes the weight of empty pycnometer, *E* represent the weight of pycnometer + distillate and *V* is the volume of pycnometer. With the density known the ethanol content can be read off standard conversion tables. Although the theory of the method is very simple, the technique is time consuming and requires an experienced technician, making it expensive and susceptible to error (Brereton *et al*., 2003). Another means of determining the ethanol concentration after distillation is with an oscillation-type density meter. This is the wet chemistry method preferred by the alcoholic beverage industry and also used in this study to determine the reference values for the calibration, as it is rapid and simple to perform. An on-board calculation on the oscillation-type vibrating-U-tube density meter (model DMA-58, Anton Paar) used in this study converts density to ethanol. The oscillation-type density meter will be discussed in full under the density heading. Pycnometry demonstrated greater variability in terms of precision than oscillation-type density meter, because of the greater opportunity for experimental error in making the weight measurements necessary for pycnometry (Lachenmeier *et al*., 2005). The prior distillation of both techniques leaves room for error, hence the setting for this study.

A hydrostatic balance is based on Archimedes' principle that the upward buoyancy force is equal to the weight of fluid displaced by the submersed body, which is related to the volume of the body and the density of the displaced fluid and subsequently the ethanol content. This technique is becoming more popular due to developments in the electronic instruments (Brereton *et al*., 2003). Although the hydrostatic balance method is gaining popularity, the sampling frequency remains low and may require hours for a single measurement (González-Rodrígues *et al*., 2003).

Ethanol determination by ebulliometry is relatively fast, easy and inexpensive (López Mahia *et al.*, 1992). However the accuracy compared to the other techniques mentioned above is poor. In addition, since high sugar levels may lower the boiling point, alcoholic beverages containing >2% residual sugar (RS) should be compensated for as well as high ethanol values (>14 %v/v) that could cause interferences, thus this method is not suitable for the determination of ethanol content in spirit products.

The enzymatic method is based on the oxidation of ethanol to acetaldehyde in the presence of NAD^+ by means of a alcohol dehydrogenase (ADH) catalysed reaction:

$$EtOH + NAD^+ \xleftrightarrow{ADH} acetaldehyde + NADH + H^+$$

The coloured NADH formed, which is directly proportional to the concentration of ethanol, is then determined by spectrophotometry (Svensson *et al*., 2005). Though widely used in experiments, the enzymatic methods involve costly measurements and are affected by the sample colour, gaseous ethanols, carbon dioxide gas. In addition, these enzymatic methods require dilution of the sample, which requires precision and careful avoidance of contamination (Nakamura *et al.*, 2009).

2.3.2 Density

A variety of density sensors are currently in use. A resonating glass or metal tube is most often used for accurate density readings (three to six digits of accuracy). In the 1960s and 1970s, electronically controlled U-shaped oscillating metal and glass tubes and temperature control were applied to the manufacturing density meters. While filled with a fluid, the tube is driven into resonance electrostatically and its motion sensed using metal electrodes placed under the microtube (Sparks *et al.,* 2003). The square of the resonance frequency is inversely proportional to the sum of the mass of tube and tube contents. As both the tube mass and tube inner volume are known values, the vibrating tube method allows the density of unknown fluids to be determined in a single measurement. The U-tube is kept oscillating continuously at the characteristic frequency, which depends on the density of the filled-in sample. The oscillation period is measured and converted into density by the equation of the Mass-Spring-Model:

$$F = \frac{1}{2}\pi\sqrt{\frac{c}{(M + \rho V)}}$$

Where *F* is the frequency, *c* indicates the spring constant, *M* is the mass, ρ the density and *V* the volume (González-Rodríguez *et al*., 2003).

2.3.3 Obscuration

Obscuration is the deviation from the actual ethanol strength due to the presence of dissolved substances in brandies. This parameter is important in the final product quality control as it gives an indication of the presence of dissolved substances in the spirits, in particular, but not only, the sugar content. Obscuration is an important specification in brandy production as it contributes greatly to the mouth-feel of the product. The obscuration expresses the ‘degree of sweetness’ of a brandy (Le Roux, 1997).

The obscuration is determined by calculating the difference between the true ethanol strength (TS) after distillation, and direct ethanol strength (DS) value, obtained directly by a density meter of the undistilled product. When a difference exists, as it does in brandies, this difference is called the obscuration of the sample. It is generally accepted that a residual sugar (RS) content of 15 g/L gives an obscuration of 3 and that there is a direct relationship between obscuration and RS (Le Roux, 1997).

2.3.4 Colour

One of the main sensory parameters for the quality of foods is their colour, and it is the first characteristic attracting consumers' attention. Thus, it is considered a major feature for the assessment of food product quality. Colour can be assessed by both visual and instrumental procedures. Generally, in the instrumental assessment of food colour, spectrophotometers are used to quantify reflectance, transmittance or absorbance characteristics (Martin *et al*., 2007). Colour measurements are part of the final quality check of a product. Colourimetry is the science of measuring and evaluating colour (Zwinkels, 1996). A colourimeter is a term used to designate an instrument for absorption measurements in which the human eye serves as the detector using one or more colour comparison standards (Skoog *et al*., 1998). Since the human eye is the detector, the results are subjective and biased to personal experience and preferences.

2.4 ANALYTICAL METHOD VALIDATION

In late 1999, the International Organisation for Standardisation (ISO) and the International Electrotechnical Commission (IEC), issued the ISO/IEC 17025 international quality standard, which incorporates all of the necessary requirements for testing and/or calibration laboratories, to prove their technical competence and validity of the data and results they produce (Vlachos *et al*., 2002, ISO, 2005). Analytical test method validation is done to ensure that an analytical methodology is accurate, specific, reproducible and robust over the specified range that an analyte will be analysed. Method validation provides an assurance of reliability during normal use (Shabir, 2003).

Step one of method validation is testing the sample set for outliers by performing the Dixon Q-test (Mermet, 2008). The hypothesis tested, is:

H_0: The distance between the suspect value and its closest neighbour is within limits (Q calculated value < Q critical value)

H_1: The distance between the suspect value and its closest neighbour is above the limits (Q calculated value > Q critical value)

The Q-value is calculated by the following equation:

$$Q_{calc} = \frac{(suspectvalue - nearestvalue)}{Range}$$

Where *Range* is the highest value in the data sequence minus the lowest value. The Q calculated value is compared to the Q critical value in the Dixon Q-table (Table 2.2):

Table 2.2: Dixon table (Miller & Miller, 1984)

	Level Of Confidence (LOC)		
Number of observations	**(90%) $Q_{0,10}$**	**(95%) $Q_{0,05}$**	**(99%) $Q_{0,01}$**
3	0.941	0.970	0.994
4	0.765	0.829	0.926
5	0.642	0.710	0.821
6	0.560	0.625	0.740
7	0.507	0.568	0.680
8	0.468	0.526	0.634
9	0.437	0.493	0.598
10	0.412	0.466	0.568
15	0.338	0.384	0.475
20	0.300	0.342	0.425
25	0.277	0.317	0.393
30	0.260	0.298	0.372

When the Q-calculated value is higher than the Q-critical value for the suspect value, H_0 is rejected, and the value is removed from further calculations. This may only be applied once to a data set, thus only one outlier may be removed, since the removal of any data affects the range of the data set.

After the sample set has been cleared of outliers, the precision is measured. This is given as the standard deviation (SD), the range and the coefficient of variation (CV) (Mermet, 2008). The SD is the root mean square of deviation from the mean of the set with *n* number of samples calculated with equation:

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

where x_i is item *i* in the set, $\bar{x}$ is the mean of the set and *n* is the number of samples.
CV is determined with equation:

$$CV = \left(\frac{SD}{Mean}\right) \times 100$$

Where the *mean* is calculated as the sum of the variable values divided by the number of samples. Several authors refer to the CV as % Relative Standard Deviation (%RSD). This used as a standard procedure to measure instrument precision and has been used in various wine quantification FT-IR studies (Soriano *et al*., 2007; La Torre *et al*., 2006; Lachenmeier, 2007). To keep to the related studies, further mention will be made to the %RSD for the CV.

The confidence interval presents the interval on the measurement scale within which the true value lies with a specified probability (i.e. 95%). Within this interval, the result is regarded as being accurate (Taverniers *et al*., 2004).

The ruggedness of a method is defined as its ability to remain unaffected by small deliberate variations in method parameters (Shabir, 2003). This criterium is evaluated by varying method parameters such as percent organic solvent, pH of buffer in mobile phase, ionic strength, etc. (Shabir, 2003). In this study ruggedness was evaluated by changing distilling points, using

different sample sizes and using different products. Analysis of variance (ANOVA) tests were used in this study to determine possible bias contribution (Jurado *et al*., 2007) and to investigate possible interactions between variables. This test proves or disproves the statistical significant differences with changes applied to the experimental lay-out. ANOVA test is further clarified in the Univariate statistics segment.

The accuracy of the methods is expressed as the Standard Error of Laboratory (SEL). SEL was calculated as in equation:

$$SEL = \sqrt{\frac{\Sigma(y_1 - y_2)^2}{2n}}$$

Where y_1 and y_2 are the values from duplicate determinations and n is the number of samples.

2.5 UNIVARIATE AND MULTIVARIATE DATA ANALYSIS

Extracting the maximum amount of information from gathered data is the aim of all chemists. Svante Wold (1995) defined chemometrics in 1994 as: 'How to get chemically relevant information out of measured chemical data, how to represent and display this information, and how to get such information into data' (Wold, 1995). Instrumentation giving multivariate responses to each sample quickened the statistic progression from univariate to multivariate analysis. The univariate and multivariate statistics that are discussed have been described in standard statistical textbooks and were used in this study (Martens & Martens, 2001; Esbensen, 2002; Manly, 2005).

2.5.1 Univariate data analysis

Univariate data analysis, also known as descriptive statistics, deals with one or two variables at a time. This gives valuable descriptive information about the data set. These methods include calculation of the average, standard deviation (SD), standard error of laboratory (SEL) and significant differences. The average of a data set of values is the sum of the values divided by the number of samples, giving an indication of the central location of the data set. The SD describes the variability in the data set, giving the typical value a sample will deviate from the average. The SEL is used to determine the measuring error of the analytical method (see method validation 2.4). Analysis of variance (ANOVA) analyses the effect of variables. This can determine significant differences and interactions between variables (Esbensen, 2002; Luciano & Næs, 2009).

2.5.2 Multivariate data analysis

Multivariate data analysis takes interactions between parameters into account. Multivariate data analysis is used for a number of different purposes, namely data exploration, regression and prediction, and discrimination and classification (Esbensen, 2002). Various techniques have branched from these three functions to multivariate data analysis. Only the techniques used in this study will be described further. In this study the **x**-variables refer to the infrared spectra gathered, while the **y**-variables refer to the reference method analysis results. The observations in a study refer to each sample.

2.5.2.1 Data exploration/description

Data exploration means taking a look at the data to find interesting phenomena (Esbensen, 2002). As a result, outliers, clustering of objects and gradients between clusters may be detected (Geladi, 2003). Outliers can be described as samples that deviate from the normal pattern in a particular data set (Esbensen, 2002). Principal Component Analysis (PCA), is one of the most commonly applied techniques in analysis of data generated in the alcoholic beverage industry (Cozzolino *et al*., 2005; Pontes *et al*., 2006 Palma & Barroso, 2002; Cozzolino *et al.*, 2007; Nordon *et al.,* 2005; Picque *et al.*, 2006; Boulet *et al.*, 2007; Ferrari *et al*., 2011) and is used for data description and explorative data structure modelling.

PCA aims to model the structure in a data set through linear combinations of the original variables, selected to maximise the variation between the samples (Esbensen, 2002). Relationship between samples can be visualised by score plots, facilitating identification of subgroups and detection of possible outliers (Bäckström *et al*., 2007; Cozzolino *et al*., 2006). Samples that share similar properties will group together on a score plot (Esbensen, 2002). The loadings plot will give valuable input of what causes the specific groupings of samples. Figure 2.1 is an example of using PCA for outlier and clustering detection (Lachenmeier, 2007).

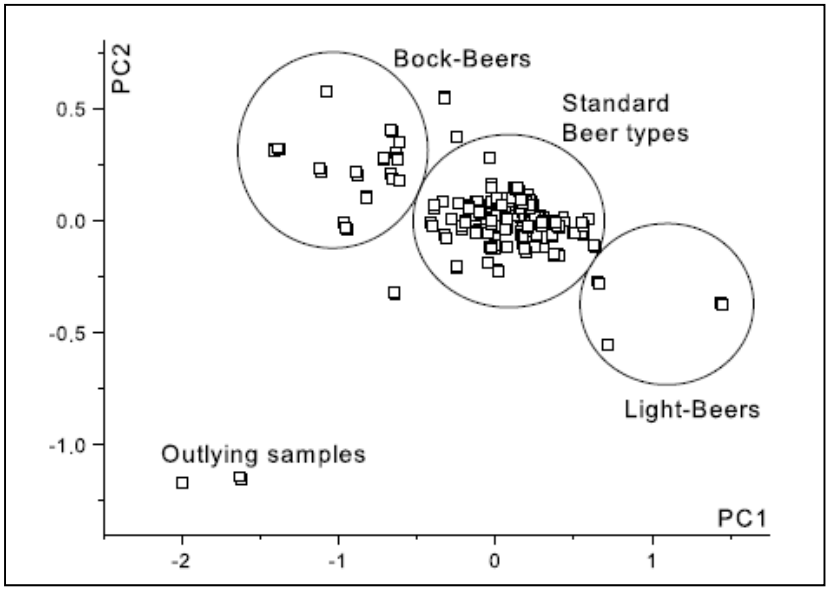


Figure 2.1: Example of utilising PCA for outlier and clustering detection (adapted from Lachenmeier, 2007).

2.5.2.2 Regression and prediction

Regression is an approach for relating two sets of variables to each other, thus determining **y**-variables (i.e. chemical concentration) from the relevant **x**-variables (i.e. spectra). Prediction means determining **y**-values for new **x**-objects, based on a calibration model, thus only relying on the new *X*-data (Esbensen, 2002).

The most popular regression method for multicomponent analysis for infrared analysis in the alcoholic beverage industry, is partial least squares (PLS) (Cozzolino *et al.,* 2009; Paul, 2009). PLS is built on PCA technology.

The optimal number of PLS factors are manually selected based on the lowest error. The statistical measurements for evaluating the calibration models included bias (calculated by equation below), root mean squared error of cross validation (RMSECV), when based on the

calibration set and root mean squared error of prediction (RMSEP) when based on the validation set.

$$\boldsymbol{bias} = \frac{\mathbf{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{y_i} - \boldsymbol{\hat{y}_i})$$

where y_i is the reference value for the *i*th sample; $\hat{y}_i$ is the predicted value for the *i*th sample; *n* is the number of samples (Næs *et al*., 2004).

RMSECV is calculated by equation:

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{c}_i - c_i)^2}{n}},$$

where $\hat{c}_i$ is the predicted concentration, c_i the actual reference concentration, *n* the number of samples used in the calibration model (Esbensen, 2002). In the assessment of the validation set, RMSECV is substituted by RMSEP, where *n* is the number of samples in the validation set (Lobo *et al*., 2006). The error achieved for the calibrations is compared to the SEL which is determined during the validation of the reference methods as mentioned in section 2.4. Correlation coefficient (R^2) indicates the precision achieved by the calibration model (Esbensen, 2002) and gives an indication of how well it may be expected to work on new samples. The residual predictive deviation (RPD) was used as a tool to evaluate the prediction ability of the calibration model. It is defined as the ration between the SD and the error for the prediction (Pink *et al*., 1998).

2.5.2.3 Discrimination and classification

Soft Independent Modeling of Class Analogy (SIMCA) is the most used class-modeling technique (Berrueta *et al.,* 2007), see Table 2.4 for examples of applications in the alcoholic beverage industry. With SIMCA, a PCA model is created for each class. SIMCA determines whether an observation belongs to a specific class based on the distance thereof to a specific model. A useful tool for the interpretation of SIMCA results is the Cooman's plot, which shows the discrimination between two classes (Berrueta *et al.,* 2007).

2.5.2.4 Outlier detection

In addition to using the PCA plots for identification of outliers, other techniques that aid the identification of outlier (samples with atypical spectra) samples are X-Y relation plots, Hotelling T^2 and Distance to model X (DModX) plots, used in this study. X-Y relation plots show the relationship between samples in the Y-space and the variables in the X-space constructed with the Unscrambler software (version 9.2, Camo ASA, Trondheim, Norway). With regards to the DMod X plots, a large value for an observation, indicates that the observation is far from the other observations in the *X*-model space. With a 95% confidence level, observations outside of the range indicate that the chance the observation belongs to the specific group is less than 5% (Simca User Guide, 2008). With Hotelling T^2 statistic, the relationships between variables are taken into account by means of the covariance matrix, which is used to weigh the relative distance between an observation and the sample mean (Cedeño Viteri *et al.,* 2012).

2.6 INFRARED SPECTROSCOPY

The success of infrared (IR) studies in the alcoholic beverage industry can be attributed to a number of reasons: speed of analysis, sensitivity, user-friendliness and versatility of sampling techniques for various forms of samples. Convenience of spectra evaluation is also an important feature (Gauglitz & Vo-Dinh, 2003). Fourier transform infrared (FT-IR) technology is a subsection of IR studies based on the range 400 to 4000 cm^{-1}, measuring absorption of chemical bonds in organic functional groups (Smith, 1999). Figure 2.2 shows an example of FT-IR spectra of beer (blue lines) and spirit drinks (black lines) (Lachenmeier, 2007). The figure shows the spectra of beer and spirits are chemically similar and display similar and overlapped absorptions. The characteristic water absorption areas are also indicated on the figure.

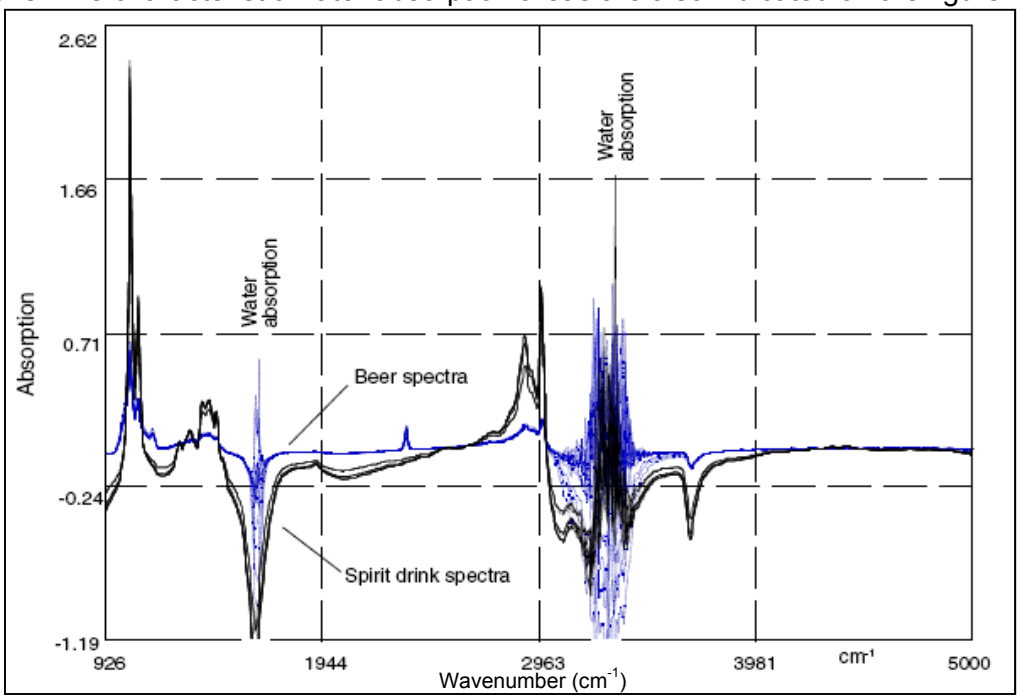


Figure 2.2: FT-IR spectra of 10 typical beer samples (blue lines) and 10 typical spirit drinks (black lines) (adapted from Lachenmeier, 2007).

The first purpose-built wine analyser based on FT-IR technology, the WineScan FT120 (Foss Analytical, Denmark), was marketed in 1998. The WineScan uses FT-IR spectroscopy together with multivariate statistics to correlate the spectral response of a sample with compositional data as determined by reference laboratory methods. The use of a FT-IR instrument with commercially available ready-to-use calibration models for different products is an advantage for unskilled users and for routine analysis.

2.6.1 A review of quantitative studies with IR spectroscopy in the alcoholic beverage industry

Quantitative analysis is the determination of the concentration of a particular substance in a sample. Compounds can be determined from spectra if a calibration model correlating the IR spectrum to the analytical reference result is obtained (Moreira & Santos, 2004). Reviewing the available literature it became apparent that IR spectroscopy is a well established and preferred method for the quantification of various parameters in wine, summarised in Table 2.3. Since

spirit products are the subject of this studies, a few interesting studies on spirit products from Table 2.3 are highlighted below.

Nordon *et al.* (2005) investigated the use of non-invasive NIR- and Raman spectrometry for non-destructive determination of ethanol content of spirits through the widest part of 700 mL static bottles. The proposed methods could be used to calculate the average ethanol concentration over a number of bottles in a bottling line, non-destructive analysis of samples in bottles in a quality control laboratory or testing for counterfeit products without opening the bottles (Nordon *et al*., 2005).

FT-IR spectroscopy combined with multivariate data analysis was used by Lachenmeier (2007) for the quality control and authenticity assessment of 535 spirit drinks and 461 beers. The reported results indicated great accuracy for the determination of spirit parameters density, ethanol, methanol, ethyl acetate, propanol-1, isobutanol and 2-/3-methyl-1-butanol (R^2=0.90-0.98), as well as beer parameters ethanol, density, original gravity and lactic acid (R^2=0.97-0.98). The results suggest that FT-IR is a useful tool in the quality control of spirit products and beer (Lachenmeier, 2007).

The determination of ethanol in all types of alcoholic beverages was further explored with on-line liquid-liquid extraction of ethanol with chloroform and FT-IR. Results suggested that samples with ethanol higher than 15 %v/v required dilution with double de-ionised water (Gallignani *et al.,* 2005).

MIR spectrometry with a diamond ATR immersion probe and polycrystalline silver halide fibres has been used for the direct and simple determination of the ethanol concentration in whisky and the identification of counterfeit samples. Univariate and multivariate calibration with an average relative error of 1.2% and 0.8%, respectively; distinguished between different caramel colourants and different whisky samples. The methodology could also be used to distinguish between authentic whiskies containing no caramel and counterfeit samples (McIntyre *et al*., 2011).

In 1994 Gallignani *et al.* (1994) proved accurate determination of ethanol in alcoholic beverages, from beer to spirit samples, by derivative FT-IR. Prior dilution of spirits was required (Gallignani *et al.*, 1994).

The successful determination of ethanol, density and total dry extract in spirits and liqueurs was reported by Arzberger and Lachenmeier (2008) by FT-IR spectroscopy and PLS regression.

Lachemeier *et al.* (2010) described a mobile flow-through infrared device for the determination of ethanol concentration in wine, beer and spirits. The methodology was developed for the labelling control of wine, beer and spirits, or the process monitoring of fermentation (Lachenmeier *et al.,* 2010)

Table 2.3: Review table of the application of IR spectroscopy for quantitative analysis in the alcoholic beverage industry.

Author, year of publication	Journal	Objective	Samples	Statistical analysis	Conclusions
Urbano-Cuadrado *et al*., 2004	Analytica Chimica Acta 527, 81-88	Evaluation of NIRS[a] to the evaluation of 16 enological parameters in wine.	180 samples red, rosé and white wines - young and aged wines of different grape varieties.	NIRS[a], PLSR[b]	Accurate determination of ethanol, volumic mass, total acidity, pH, glycerol, colour, tonality and total polyphenol index. Screening capabilities for volatile acidity, organic acids, reducing sugars and total sulphur dioxide.
Patz *et al*., 2004	Analytica Chimica Acta 513, 81-89	Evaluation of FT-MIR[c] for the determination of 19 parameters in wine.	327 German wines	FT-MIR[c], PLSR[b]	Excellent quantitative results were obtained for: ethanol, relative density, extract, conductivity, glycerol, total phenol, Trolox equivalent antioxidative capacity, fructose, glucose, sugar and total acid. Calibration model could be transferred between FT-MIR[c] machines with the same hardware.
Kupina & Shrikhande, 2003	American Journal of Enology and Viticulture, 54, 131-134	Evaluation of FT-IR[d] for quality control wine analyses	256 wines of 5 different types	FT-IR[d], PLSR[b]	Good correlation obtained for parameters ethanol, titratable acidity, pH, volatile acidity and reducing sugars.
Gishen & Holdstock, 2000	The Australian Wine Research Institute Annual Technical Issue, 75-80	Evaluation of the WineScan for application in routine wine analysis for ethanol, glucose/fructose, pH, total acid and volatile acid.	173 different types of wines	FT-IR[d], multivariate calibrations	Good correlation with the reference laboratory methods for ethanol, pH, titratable acidity and volatile acidity was achieved. Glucose/fructose was determined reasonably well. Results showed the calibration for the prediction of total sulphur dioxide was suitable for screening.
Vonach *et al*., 1998	Journal of Chromatography A 824, 159-167	Evaluation of an advanced flow cell HPLC-FTIR[e] for direct determination of components in wine.	3 red & 3 white Austrian wines	HPLC–FTIR[e], Bruker 3-dimensional data treatment	The capability of real time HPLC–FTIR[e] for the determination of carbohydrates, alcohols and organic acids in wines was demonstrated for the time.

Author, year of publication	Journal	Objective	Samples	Statistical analysis	Conclusions
Urbano-Cuadrado *et al*., 2005	Talanta 66, 218-224	The applicability of the spectroscopic techniques in the near and mid infrared zones to the determination of wine parameters.	180 samples of different varieties and origins	NIRS[a] & FT-MIR[c], PLSR[b]	NIRS results were better than those obtained by FT-MIR[c] due to the high signal/noise ratio of the latter. The combination of both spectral zones has been studied for the first time. The equations for each zone can only be used for screening
Nieuwoudt *et al*., 2006	Journal of Microbiological Methods 67, 248-256	Screening of fermentation profiles of a selection of glycerol-overproducing *Saccharomyces cerevisiae* wine yeasts strains.	Chenin blanc & synthetic musts	FT-IR[d], PLSR[b], PCA[f]	Excellent quantitative prospects for parameters: volatile acidity, ethanol, glycerol and residual sugar for the Chenin blanc.
Urtubia *et al*., 2004	Talanta 64, 778–784	Development of infrared calibrations for monitoring glucose, fructose, glycerol, ethanol and organic acids during large scale wine fermentations of Cabernet Sauvignon.	273 samples from large scale fermentation tanks	FT-IR[d] & MIR[i], PLSR[b]	Developed calibrations provided good estimations for glucose, fructose, organic acids, glycerol and ethanol during the entire fermentation of Cabernet Sauvignon musts. Distinctions could be made between a normal and a problematic fermentation.
Soriano *et al*., 2007	Food Chemistry 104, 1295-1303	Feasibility of FT-IR[d] spectroscopy for determination of anthocyanins in red wines.	350 samples of young red wines	FT-IR[d], PLSR[b]	WineScan FT 120 analyser is suitable for routine laboratory measurement of anthocyanins and provides additional information regarding red wine colour.
Romera-Fernández *et al*., 2012	Talanta 88, 303-310	Feasibility study of using FT-MIR[c] combined with chemometrics for the determination of anthocyanins in red wines of different degrees of ageing	158 red wines from 11 wineries	FT-MIR[c], PCA[f] & PLSR[b]	FT-MIR[c] instrument calibration is a useful tool for a quick determination of the anthocyanin content of young wines of the current vintage.

Author, year of publication	Journal	Objective	Samples	Statistical analysis	Conclusions
Cozzolino *et al*., 2008	Talanta 74, 711–716	The use of VIS[g] and NIRS[a] to measure the concentration of elements in Australian wines was investigated.	32 white and 94 red wine samples	VIS[g] & NIRS[a],+ PLSR[b]	Relationships exist between NIRS[a] spectra and the concentration of some elements in wine.
Schneider *et al*., 2004	Analytica Chimica Acta 513, 91–96	Development of a new method using FT-IR[d] spectrometry and chemometric techniques to determine glycosidic precursors.	39 samples	FT-IR[d], PLSR[b]	This new method allows nine samples in 2 days to be analysed versus 5 days for the reference method.
Cozzolino *et al*., 2006	Analytica Chimica Acta 563, 319-324	Assess the feasibility of combining MS-eNose[h] and VIS[g] + NIRS[a], coupled with chemometrics, to predict the sensory scores in commercial Riesling wines grown in Australia.	20 commercial Australian Riesling wines	MS-eNose[h], VIS[g] and NIRS[a], PLSR[b]	The results suggested that data from instrumental techniques coupled with chemometrics might be related with sensory scores measured by a trained panel.
Nordon *et al.*, 2005	Analytica Chimica Acta 548, 148-158	Evaluation of non-invasive NIRS[a] and Raman spectrometries for determination of ethanol content in spirit products.	32 samples of whiskies, vodkas and alcoholic sugary drinks	NIRS[a] & Raman spectrometries, PCA[f], PLSR[b]	Non-invasive measurements could be used for the non-destructive analysis of samples in bottles in a quality control laboratory.
Lachenmeier, 2007	Food Chemistry 101, 825-832	Evaluation of FT-IR[d] in combination with PLSR[b] as a complete multi-component screening method for spirit drinks and beer quality control and authenticity assessment in official food control.	535 spirit drinks and 461 beers	FT-IR[d] + PLSR[b], PCA[f]	Great accuracy demonstrated for spirit parameters density, ethanol, methanol, ethyl acetate, propanol-1, isobutanol and 2-/3-methyl-1-butanol ($^{l}R^2$=0.9-0.98), as well as for beer parameters ethanol, density, original gravity and lactic acid ($^{l}R^2$=0.97-0.98). Differentiation of deteriorated fruit spirits distilled from microbiologically spoiled mashes was possible. PCA[f] classification for authenticity control is possible.

Author, year of publication	Journal	Objective	Samples	Statistical analysis	Conclusions
Gallignani *et al*., 2005	Talanta 68, 470-479	Design a flow injection FT-IR[d] spectrometric procedure for the direct determination of ethanol in all types of alcoholic beverages	Commercial alcoholic beverages, from beers to spirits	FT-IR[d], multivariate calibrations	Samples with ethanol higher than 15%v/v required dilution. This methodology represents a valid alternative for the determination of ethanol in alcoholic beverages, and could be suitable for the routine control analysis.
Iñón *et al*., 2006	Analytica Chimica Acta 571, 167–174	Evaluation of a methodology based upon NIRS[a] and MIRS[i] and chemometric data treatment for the evaluation of beer parameters.	43 samples of different types of beer	NIRS[a], MIRS[i], ANN[j] & PLSR[b]	The determination of real extract, original extract and ethanol in beers can be successfully carried out through the combination of MIRS[i] and NIRS[a] techniques. ANN[j] obtained better predictive capabilities than PLSR[b].
Llario *et al*., 2006	Talanta 69, 469–480	Evaluation of ATR-FTIR[k] with PLSR[b] to estimate quality parameters in beer.	45 samples of different types of beer	ATR-FTIR[k], PLSR[b] & cluster analysis	Relative prediction errors of 1.5% for ethanol, 2.8% for real extract and 1.9% for original extract was achieved. Classification of samples, from the MIRS[i] was achieved; related to the original extract and ethanol content. It is supposed that clustering of most similar samples may be due to the type and content of different carbohydrates in the beer.
Gallignani *et al*., 1994	Analytica Chimica Acta 287, 275-283	Development of a procedure for the on-line derivative FT-IR[d] measurements of ethanol in all types of alcoholic beverages.	Beer, wine, vodka, gin and whisky	Derivative FT-IR[d]	Methodology provides accurate results in determination of ethanol in alcoholic beverages. Dilution of spirits required

Author, year of publication	Journal	Objective	Samples	Statistical analysis	Conclusions
Lobo *et al*., 2006	LWT-Food Science & Technology 39, 1026-1032	Developing and validation of prediction models for the routine analysis of cider.	Ciders exclusively made from cider-apple pressing, at different stages of the making process, from the end of the fermentation to several months in bottle.	FT-IR[d], PLSR[b]	Reliable and suitable calibration models were optimised for the routine analysis of specific gravity, total acidity, volatile acidity, pH, ethanol and fructose in ciders.
Arzberger & Lachenmeier, 2008	Food Analytical Methods, 1, 18-22	Characterising ethanol strength, density, and total dry extract in spirits and liqueurs	163 liqueurs, 298 spirits	FT-IR[d] spectroscopy, PLSR[b]	FT-IR[d] spectroscopy and PLSR[b] is well suited for the analysis of density ethanol and total dry extract in spirits and liqueurs.
Lachenmeier *et al.*, 2010	Chemistry Central Journal, 4, 1-13	Mobile determination of ethanol strength in wine, beer and spirits using a flow-through infrared sensor.	Commercial wines, spirits and beers	Infrared sensor + flow through cell, PLSR[b]	Description of a mobile flow-through infrared device to determine ethanol strength.
Nieuwoudt *et al.*, 2004	Journal of Agricultural and Food Chemistry, 52, 3726 - 3735	Determination of glycerol in wine	329 wines of various styles	FT-IR[d], PCA[f] & SIMCA[k]	Calibration sets have to be carefully selected in order to design calibration models that find a balance between robustness and accuracy of prediction.
Regmi *et al.*, 2012	Analytica Chimica Acta 732, 137 – 144	Determination of organic acids in wine and wine-derived products	155 brandy, 138 white wine, 124 red wine, 17 sweet wine and 57 commercial vinegar samples	FT-IR[d], PLSR[b]	Calibration of the FT-IR[d] spectroscopy method depends very strongly on the composition of the sample set and on the quality of the reference analysis.

Abbreviations used:

[a]NIRS – Near infrared spectroscopy; [b]PLSR – Partial least squares regression; [c]FT-MIR – Fourier transform mid infrared; [d]FT-IR – Fourier transform infrared; [e]HPLC-FTIR - high-performance liquid chromatography – Fourier transform infrared; [f]PCA – Principal component analysis; [g]VIS – visible spectroscopy; [h]MS-eNose – Mass Spectrometry electronic nose; [i]MIRS – Mid Infrared Spectroscopy; [j]ANN – artificial neural networks; [k]ATR-FTIR – attenuated total reflectance-Fourier transform infrared; [l]R^2 – correlation coefficient

2.6.2 A review of qualitative studies with IR spectroscopy in the alcoholic beverage industry

Other applications of IR spectroscopy do not involve the specific quantification of a parameter, or confirmation of the chemical structure of a substance, instead, a qualitative test is required for the presence of a substance or to differentiate between groupings of samples. The foundation of this is that the IR spectrum functions as a unique chemical fingerprint of a sample (Cozzolino *et al.,* 2011; Bevin *et al.,* 2008). Table 2.4 summarises the qualitative studies of infrared spectroscopy in the alcoholic beverage industry. In general, the goals of these studies were to develop models to classify samples according to different criteria, such as country of origin, product type, detection of adulteration, etc.

Traditionally techniques used to verify authentication of products/brands, determine certain marker compounds and compare these results with samples verified as authentic. These techniques are often laborious and time consuming and the amount of parameters that needs to be quantified for a correct classification increases with more innovative ways of counterfeiting products. Spectroscopy gives a holistic view of authenticity testing, viewing the chemical composition of a sample in the whole. Brand authentication in brandies (Sádecká *et al.,* 2009), tequilas (Barbosa-García *et al.,* 2007; Contreras *et al*., 2010) and beer (Engel *et al.,* 2012) have been investigated using spectroscopic techniques with chemomentric analysis.

Classifying wine or spirit products into types or groupings, such as origin, production style and age by spectroscopic techniques (Cozzolino *et al.,* 2009; Mignani *et al.,* 2012; Palma & Barosso, 2002; Di Egidio *et al.,* 2011; Picque *et al.,* 2006) have also grabbed the attention of researchers.

Adulteration of alcoholic beverages is a concern worldwide. The main risk for consumers is the ingestion of illegal raw materials, especially methanol (Pontes *et al.,* 2006). Ferrari *et al.* (2011) presented a study to discriminate wines containing anthocyanins originated from black rice to elaborate the colour in red wine and grapevine by using spectroscopic techniques. Pontes *et al.* (2006) proposed a methodology using Near-infrared (NIR) spectrometry and chemometric methods to classify and verify whiskey, brandy, rum and vodka samples. The proposed method was successfully applied in the verification of alcoholic beverages adulteration with 100% of correct prediction at 95% of confidence level (Pontes *et al*., 2006).

Table 2.4: Review table of the application of IR spectroscopy for qualitative analysis in the alcoholic beverage industry.

Author, year of publication	Journal	Objective	Samples	Statistical analysis	Conclusions
Boulet *et al*., 2007	Carbohydrate Polymers 69, 79-85	Characterising the spectra of various purified polysaccharides.	Study done in France	FT-IR[a], PCA[b]	Significant differences found between main polysaccharide families- Rhamnogalacturonans (RG-Is, RG-Iis), Polysaccharide rich in Arabinose and Galactose (AGPs) and Mannoproteins (MPs)
Cozzolino *et al*., 2009	Food Chemistry 116, 761–765	Classify organic and non-organic wines produced in Australia.	172 commercial wine samples	MIR[c], PLSR[d], PCA[b] & LDA[e]	The LDA[e] models based on the PCA[b] scores correctly classified on average, more than 75% of the wine samples while the DPLS[f] models correctly classified more than 85% of the wines belonging to Organic and Non-Organic production systems, respectively. MIR[c] combined with discriminant techniques might be a suitable method to classify wines produced under organic systems.
Bevin *et al*., 2008	Analytica Chimica Acta 621, 19–23	Discrimination between different red and white wine varieties	119 red and 72 white wine samples were collected from HardyWine Company wineries across different regions of Australia	MIR[c], PCA[b] & LDA[e]	The technique predicted the variety with greater than 95% success on an external dataset. Overall, the results of this study suggested that this technique can therefore be used as a robust rapid screening tool for varietal discrimination by the wine industry.
Pontes *et al*., 2006	Food Research International 39, 182-189	Classification and verification of adulteration of distilled alcoholic beverages using NIR[g] spectroscopy and methods of chemometric classification.	69 samples of Whiskey, brandy, rum and vodka	NIR[g], PCA[b] & SIMCA[h]	Pure and adulterated samples were successfully classified – 100% at 95% confidence level

Author, year of publication	Journal	Objective	Samples	Statistical analysis	Conclusions
Ferrari *et al*., 2011	Analytica Chimica Acta 701, 139– 151	Discrimination of wines containing anthocyanins originated from black rice and grapevine.	35 wine samples	NIR[g], PLSR[d]-discriminant analysis	Satisfactory results were obtained on NMR[i] spectra in the aromatic region between 6.5 and 9.5 ppm. The classification method based on wavelet-based variables selection, permitted to reach an efficiency in validation greater than 95%.
Palma & Barroso, 2002	Talanta 58, 265-271	Differentiation and classification of wines and brandies during their ageing process, as well as for the characterisation and differentiation of distilled drinks from several producing countries	Fino sherries (18), Jerez brandies(59) & distilled drinks(33) including Spanish brandies, French brandies, cognacs, armagnacs and South African brandies	FT-IR[a], PCA[b], LDA[e] & PLSR[d]	PLSR[d] obtained a regression between the degree of ageing of brandies of Jerez and the data of the FT-IR[a] spectrum (correlation coefficient=0.986). Spanish, French and South African brandies, as well as cognacs and armagnacs have been characterised, and a complete differentiation of the latter two types from the rest of the samples of distilled drinks has been obtained.
Picque *et al*., 2006	Journal of Agriculture and Food Chemistry 54, 5220-5226	Discrimination between Cognacs and other distilled drinks such as whiskies, rums, brandies, Armagnacs, bourbons, and counterfeit products.	151 samples – Cognacs, Armagnacs, whiskies, brandies, bourbons and rums	MIR[c], PCA[b] & PLSR[d]-discriminant analysis	Dry extract data used to differentiate between Cognacs and Armagnacs, whiskies, bourbons and rums. And polyphenol concentration to separate Cognacs and brandies from counterfeit products. 96% of samples correctly assigned to Cognacs and non-Cognacs by PLSR[d]-discriminant analysis
Di Egidio *et al*., 2011	Food Research International 44, 544–549	Confirm the declared identity of a particular Trappist beer, namely Rochefort 8°.	275 Belgian & European beers	NIR[g], SIMCA[h], POTFUN[j]; UNEQ[k], PLSR[d]-discriminant analysis	Moderate levels of sensitivity & specificity were achieved for the characterisation of Trappist beers, with SIMCA[h] performing best. Discriminant analysis for differentiating Rochefort 8° and Rochefort 10° beers provided valuable results.
Zhang *et al.*, 2010	Journal of Molecular Structure, 974, 144-150	Discrimination of different red wine	120 samples of different kinds of wines	FT-IR[a], 2-dimensional infrared correlation spectroscopy	Infrared spectroscopy is a direct and effective method for the analysis of different red wines and discrimination thereof

Author, year of publication	Journal	Objective	Samples	Statistical analysis	Conclusions
Cozzolino *et al.*, 2011	Food Chemistry, 126, 673 - 678	Geographic classification of Sauvignon blanc wines from Australia and New Zealand.	64 wines	UV[l], visible, NIR[g], MIR[c] spectroscopy, SIMCA[h], PLSR[d]-discriminant analysis	Potential demonstrated of spectroscopy and chemometrics to classify Sauvignon blanc wines according to their geographical origin
Yucesoy & Ozen, 2013	Food Chemistry, 141, 1461 - 1465	Identification of adulterated distilled spirit, raki	34 samples, Turkish distilled spirit, raki	MIR[c], PCA[b]	Pure and adulterated samples could be separated with PCA[b]. A minimum of 0.5% methanol adulteration was successfully detected.
McIntyre *et al*., 2011	Analytica Chimica Acta 690, 228–233	Detection of counterfeit Scotch whisky samples	17 authentic and counterfeit samples of one brand of blended whisky	MIR[c], PLSR[d]	Successful categorisation of authentic and counterfeit whisky samples

Abbreviations used:

[a]FT-IR – Fourier transform infrared; [b]PCA – Principal component analysis; [c]MIR – Mid infrared; [d]PLSR – Partial least squares regression; [e]LDA – Linear discriminant analysis; [f]DPLS - discriminant partial least squares; [g]NIR – Near infrared spectroscopy; [h]SIMCA - soft independent modelling of class analogy; [i]NMR – nuclear magnetic resonance; [j]POTFUN – potential functions techniques; [k]UNEQ – unequal dispersed classes; [l]UV – ultra violet

2.7 CONCLUSION

South Arica has become a force to be reckoned with in the spirits industry, taking top honours in international competitions. In order to compete on an international level, South African distilled beverage producers have to keep up with the innovations and new technologies that greatly benefit producers in terms of quality control, authentication and economic well-being. The review on infrared spectroscopy in the alcoholic beverage industry proves that this technology has been tried and tested on various alcoholic beverages measuring various parameters as well the ability of classification application. Spirit products were introduced to this technology and vastly modified calibrations were needed to encompass the new interferences these products introduced. Multivariate data analysis has opened a gateway to multivariate information processing. Increasing production while maintaining quality is possible!

2.8 REFERENCES

1. Arzberger, U. & Lachenmeier, D.W. (2008). Fourier Transform Infrared Spectroscopy with Multivariate Analysis as a Novel Method for Characterizing Alcoholic Strength, Density, and Total Dry Extract in Spirits and Liqueurs. *Food Analytical Methods* **1**, 18 – 22.
2. Bäckström, D., Moberg, M., Sjöberg, P.J.R., Bergquist, J. & Danielsson, R. (2007). Multivariate comparison between peptide mass fingerprints obtained by liquid chromatography–electrospray ionization-mass spectrometry with different trypsin digestion procedures. *Journal of Chromatography A* **1171,** 69–79.
3. Barbosa-García, O., Ramos-Ortíz, G., Maldonado, J.L., Pichardo-Molina, J.L., Meneses-Nava, M.A., Landgrave, J.E.A. & Cervantes-Martínez,J. (2007). UV-vis absorption spectroscopy and multivariate analysis as a method to discriminate tequila. *Spectrochimica Acta Part A* **66,** 129-134.
4. Berrueta, L.A., Alonso-Salces, R.M. & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A* **1158**, 196-214.
5. Bevin, C.J., Dambergs, R.G., Fergusson, A.J. & Cozzolino, D. (2008). Varietal discrimination of Australian wines by means of mid-infrared spectroscopy and multivariate analysis. *Analytica Chimica Acta* **621,** 19-23.
6. Boulet, J.C., Williams, P. & Doco, T. (2007). A Fourier transform infrared spectroscopy study of wine polysaccharides. *Carbohydrate Polymers* **69,** 79-85.
7. Brereton, P., Hasnip, S., Bertrand, A., Wittkowski, R. & Guillou, C. (2003). Analytical methods for the determination of spirit drinks. *Trends in Analytical Chemistry* **22,** 19-25.
8. Cedeño Viteri, M.V., Rodríguez Aguilar, L.P.F., Alvarez Medina, C.R. & Sánchez, M.C. (2012). A new approach to estimate variable contributions to Hotelling's statistic. *Chemometrics and Intelligent Laboratory Systems* **118**, 12 – 126.
9. Contreras, U., Barbosa-García, O., Pichardo-Molina, J.L., Ramos-Ortíz, G., Maldonado, J.L., Meneses-Nava, M.A., Ornelas-Soto, N.E. & López-de-Alba, P.L. (2010). Screening method for identification of adulterate and fake tequilas by using UV-VIS spectroscopy and chemometrics. *Food Research International* **43,** 2356-2362.
10. Cozzolino, D., Holdstock, M., Dambergs, R.G., Cynkar, W.U. & Smith, P.A. (2009). Mid infrared spectroscopy and multivariate analysis: A tool to discriminate between organic and non-organic wines grown in Australia. *Food Chemistry* **116,** 761-765.
11. Cozzolino, D., Kwiatkowski, M.J., Dambergs, R.G., Cynkar, W.U., Janik, L.J, Skouroumounis, G. & Gishen, M. (2008). Analysis of elements in wine using near infrared spectroscopy and partial least squares regression. *Talanta* **74**, 711–716.
12. Cozzolino, D., Liu, L., Cynkar, W.U., Dambergs, R.G., Janik, L., Colby, C.B. & Gishen, M. (2007). Effect of temperature variation on the visible and near infrared spectra of wine and the consequences on the partial least square calibrations developed to measure chemical composition. *Analytica Chimica Acta* **588,** 224-230.
13. Cozzolino, D., Parker, M., Dambergs, R.G., Herderich, M. & Gishen, M. (2006). Chemometrics and visible - near infrared spectroscopic monitoring of red wine fermentation in a pilot scale. *Biotechnology and bioengineering* **95,** 1101-1107.
14. Cozzolino, D., Smyth, H.E., Cynkar, W.U., Dambergs, R.G. & Gishen, M. (2005). Usefulness of chemometrics and mass spectrometry-based electronic nose to classify Australian white wines by their varietal origin. *Talanta* **68,** 382-387.
15. Cozzolino, D., Cynkar, W.U., Shah, N. & Smith, P.A. (2011). Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New Zealand? *Food Chemistry* **126**, 673 – 678.

16. Di Egidio, V., Oliveri, P., Woodcock, T. & Downey, G. (2011). Confirmation of brand identity in foods by near infrared transflectance spectroscopy using classification and class-modelling chemometric techniques — The example of a Belgian beer. *Food Research International* **44,** 544-549.
17. Engel, J., Blanchet, L., Buydens, L.M.C. & Downey, G. (2012). Confirmation of brand identity of a Trappist beer by mid-infrared spectroscopy coupled with multivariate data analysis. *Talanta* **99,** 426-432.
18. Esbensen, K.H. (2002). Multivariate Data Analysis – In Practice. An introduction to multivariate data analysis and experimental design, *5th ed*. Camo ASA, Oslo.
19. Ferrari, E., Foca, G., Vignali, M., Tassi, L. & Ulrici, A. (2011). Adulteration of the anthocyanin content of red wines: Perspectives for authentication by Fourier Transform-Near InfraRed and ^{1}H NMR spectroscopies. *Analytica Chimica Acta* **701,** 139-151.
20. Foss Analytical, Denmark. http//www.foss.dk.
21. Gallignani, M., Ayala, C., Brunetto, M.R., Burguera, J.L. & Burguera, M. (2005). A simple strategy for determining ethanol in all types of alcoholic beverages based on its on-line liquid-liquid extraction with chloroform, using a flow injection system and Fourier transform infrared spectrometric detection in the mid-IR. *Talanta* **68,** 470-479.
22. Gallignani, M., Garrigues, S. & De la Guardia, M. (1994). Derivative Fourier transform infrared spectrometric determination of ethanol in alcoholic beverages. *Analytica Chimica Acta* **287,** 275-283.
23. Gauglitz, G. & Vo-Dinh, T. (2003). Handbook of Spectroscopy. Wiley-VCH, Weinheim, Germany. Volume 1.
24. Geladi, P. (2003). Chemometrics in spectroscopy. Part 1. Classical chemometrics. *Spectrochimica Acta Part B* **58,** 767-782.
25. Gishen, M. & Holdstock, M. (2000). Preliminary evaluation of the performance of the Foss WineScan FT120 instrument for the simultaneous determination of several wine analyses. *The Australian Wine Research Institute Annual Technical Issue* 75-80.
26. González-Rodríguez, J., Pérez-Juan, P. & Luque de Castro, M.D. (2003). Determination of ethanol in beverages by flow injection, pervaporation and density measurements. *Talanta* **59,** 691-696.
27. International Organization for Standardization (ISO), ISO/IEC 17025. (2005). General requirements for the competence of testing and calibration laboratories, ISO, Geneva, Switserland.
28. IWSC. (2012). International Wine and Spirits Competition. [URL. http://www.iwsc.net/home]. Cited the 09th of December 2012.
29. Iñón, F.A., Garrigues, S. & De la Guardia, M. (2006). Combination of mid- and near-infrared spectroscopy for the determination of the quality properties of beers. *Analytica Chimica Acta* **571,** 167-174.
30. Jurado, J.M., Martín, M.J., Pablos, F., Moreda-Piñeiro, A. & Bermejo-Barrera, P. (2007). Direct determination of copper, lead and cadmium in aniseed spirits by electrothermal atomic absorption spectrometry. *Food Chemistry,* **101**, 1296–1304.
31. Kupina, S.A. & Shrikhande, A.J. (2003). Evaluation of a Fourier Transform Infrared Instrument for Rapid Quality-Control Wine Analyses. *American Journal of Enology and Viticulture*, **54**, 131-134.
32. La Torre, G.L., Saitta, M., Vilasi, F., Pellicanò, T. & Dugo, G. (2006). Direct determination of phenolic compounds in Sicilian wines by liquid chromatography with PDA and MS detection. *Food Chemistry* **94,** 640-650.
33. Lachenmeier, D.W. (2007). Rapid quality control of spirit drinks and beer using multivariate data analysis of Fourier transform infrared spectra. *Food Chemistry* **101,** 825-832.

34. Lachenmeier, D.W., Burri, P.A., Fauser, T., Frank, W. & Walch, S.G. (2005). Rapid determination of alcoholic strength of egg liqueur using steam distillation and oscillation-type densimetry with peristaltic pumping. *Analytica Chimica Acta* **537**, 377–384.
35. Lachenmeier, D.W., Godelmann, R., Steiner, M, Ansay, B., Weigel, J. & Krieg, G. (2010). Rapid and mobile determination of alcoholic strength in wine, beer and spirits using a flow-through infrared sensor. *Chemistry Central Journal* **4:5** [http://journal.chemistrycentral.com/content/4/1/5].
36. Llario, R., Iñon, F.A., Garrigues, S. & De la Guardia, M. (2006). Determination of quality parameters of beers by the use of attenuated total reflectance-Fourier transform infrared spectroscopy. *Talanta* **69,** 469-480.
37. Lobo, A.P., Valles, B.S., Tascón, N.F., Madrera R.R. & García, O.F. (2006). Calibration models for routine analysis of cider by mid-infrared spectroscopy. *LWT - Food Science and Technology* **39,** 1026-1032.
38. López Mahia, P., Simal Gándara, J. & Paseiro Losada, P. (1992). Validation of ethanol determination in alcoholic beverages by infrared spectrophotometry using orthogonal and derivative functions to correct for water absorption. *Vibrational Spectroscopy* **3,** 133-138.
39. Luciano, G. & Næs, T. (2009). Interpreting sensory data by combining principal component analysis and analysis of variance. *Food Quality and Preference* **20,** 167-175.
40. Manly, B.F.J. (2005). The material of multivariate analysis. In: Multivariate statistical methods. A primer. 3rd ed. Chapman & Hall/CRC Publisher, New York.
41. Martens, H. and Martens, M. (2001). Why multivariate data analysis? In: Multivariate analysis of quality. An Introduction. Wiley & Sons LTD Publishers, New York.
42. Martin M.L.G.-M., Ji, W., Luo R., Hutchings J. & Heredia F.J. (2007). Measuring colour appearance of red wines. *Food Quality and Preference* **18,** 862–871.
43. McIntyre, A.C., Bilyk, M.L., Nordon, N., Colquhoun, G. & Littlejohn, D. (2011). Detection of counterfeit Scotch whisky samples using mid-infrared spectrometry with an attenuated total reflectance probe incorporating polycrystalline silver halide fibres. *Analytica Chimica Acta* **690**, 228–233.
44. Mermet, J.-M. (2008). Limit of quantitation in atomic spectrometry: An unambiguous concept? *Spectrochimica Acta Part B* **63,** 166-182.
45. Mignani, A.G., Ciaccheri, L., Gordillo, B., Mencaglia, A.A., Conzález-Miret, M.L., Heredia, F.J. & Culshaw, B. (2012). Identifying the production region of single-malt Scotch whiskies using optical spectroscopy and pattern recognition techniques. *Sensors and Actuators B: Chemical* **171-172,** 458-462.
46. Miller, J.C. & Miller, J.N. (1984). Statistics for Analytical Chemistry. *1st ed*. Ellis Horwood Limited. England.
47. Moreira, J.L. & Santos, L. (2004). Spectroscopic interferences in Fourier transform infrared wine analysis. *Analytica Chimica Acta* **513,** 263-268.
48. Næs, T., Isaksson, T., Fearn, T. & Davies, T. (2004). *A User-Friendly Guide to Multivariate Calibration and Classification.* NIR Publications, Chichester, United Kingdom.
49. Nakamura, H., Tanaka, R., Suzuki, K., Yataka, M. & Mogi, Y. (2009). A direct determination method for ethanol concentrations in alcoholic beverages employing a eukaryote double-mediator system. *Food Chemistry* **117,** 509-513.
50. Nieuwoudt, H.H., Pretorius, I.S., Bauer, F.F., Nel, D.G. & Prior, B.A. (2006). Rapid screening of the fermentation profiles of wine yeasts by Fourier transform infrared spectroscopy. *Journal of Microbiological Methods* **67,** 248-256.
51. Nieuwoudt, H.H., Prior, B.A., Pretorius, I.S., Manley, M. & Bauer, F.F. (2004). Principal Component Analysis Applied to Fourier Transform Infrared Spectroscopy for the Design of

Calibration Sets for Glycerol Prediction Models in Wine and for the Detection and Classification of Outlier Samples. *Journal of Agricultural and Food Chemistry* **52**, 3726 – 3735.

52. Nordon, A., Mills, A., Burn, R.T., Cusick, F.M. & Littlejohn, D. (2005). Comparison of non-invasive NIR and Raman spectrometries for determination of alcohol content of spirits. *Analytica Chimica Acta* **548,** 148-158.
53. Palma, M. & Barroso, C.G. (2002). Application of FT-IR spectroscopy to the characterisation and classification of wines, brandies and other distilled drinks. *Talanta* **58,** 265-271.
54. Patz, C.-D., Blieke, A., Ristow, R. & Dietrich, H. (2004). Application of FT-MIR spectrometry in wine analysis. *Analytica Chimica Acta* **513**, 81-89.
55. Paul, S.O. (2009). Chemometrics in South Africa and the development of the South African Chemometrics Society. *Chemometrics and Intelligent Laboratory Systems* **97**, 104 – 109.
56. Picque, D., Lieben, P., Cantagrei, R., Lablanquie, O. & Snakkers, G. (2006). Discrimination of Cognacs and other distilled drinks by mid-infrared spectroscopy. *Journal of Agricultural and Food Chemistry* **54,** 5220-5226.
57. Pink, J., Naczk, M. & Pink, D. (1998). Evaluation of the Quality of Frozen Minced Red Hake: Use of Fourier Transform Infrared Spectroscopy. *Journal of Agriculture and Food Chemistry*, **46** (9) 3667-3672.
58. Pontes, M.J.C, Santos, S.R.B., Araújo, M.C.U., Almeida, L.F., Lima, R.A.C., Gaião, E.N. & Souto, U.T.C.P. (2006). Classification of distilled alcoholic beverages and verification of adulteration by near infrared spectrometry. *Food Research International* **39,** 182-189.
59. Punt, C. (2010). *Economic Contribution of the Spirits and Ready-To-Drink Industries of South Africa.* Unpublished research report. University of Stellenbosch, September.
60. Regmi, U., Palma, M. & Barroso, C.G. (2012). Direct determination of organic acids in wine and wine-derived products by Fourier transform infrared (FT-IR0 spectroscopy and chemometric techniques. *Analytica Chimica Acta* **732**, 137 – 144.
61. Romera-Fernández, M., Berrueta, L.A., Garmón-Lobato, S., Gallo, B., Vicente, F. & Moreda, J.M. (2012). Feasibility study of FT-MIR spectroscopy and PLS-R for the fast determination of anthocyanins in wine. *Talanta* **88***,* 303-310.
62. Sádecká, J., Tóthová, J. & Májek, P. (2009). Classification of brandies and wine distillates using front face fluorescence spectroscopy. *Food Chemistry* **117**, 491-498.
63. South African National Standard. SANS 1841. (2008). Control of the quantity of contents in prepacked packages within the prescriptions of legal metrology legislation.
64. Schneider, R., Charrier, F., Moutounet, M. & Baumes, R. (2004). Rapid analysis of grape aroma glycoconjugates using Fourier-transform infrared spectrometry and chemometric techniques. *Analytica Chimica Acta* **513,** 91-96.
65. Shabir, G.A. (2003). Validation of high-performance liquid chromatography methods for pharmaceutical analysis. Understanding the differences and similarities between validation requirements of the US Food and Drug Administration, the US Pharmacopeia and the International Conference on Harmonization. *Journal of Chromatography A* **987,** 57–66.
66. SIMCA. SIMCA-P+ 12 User Guide. UMETRICS, Umeå, Sweden; (2008).
67. Skoog, D.A., Holler, F.J. & Nieman, T.A. (1998). *Principles of Instrumental Analysis*. 5th Edition, Harcourt Brace Company, Florida, USA.
68. Smith, B. (1999). Infrared Spectral Interpretation: A Systematic Approach. CRC Press, Boca Raton, Florida.
69. Soriano, A., Pèrez-Juan, P.M., Vicario, A., González, J.M. & Pèrez-Coello, M.S. (2007). Determination of anthocyanins in red wine using newly developed method based on Fourier transform infrared spectroscopy. *Food Chemistry* **104,** 1295-1303.
70. South African Liquor Products Act No. 60 of 1989.

71. Sparks, D., Smith, R., Schneider, R., Cripe, J., Massoud-Ansari, S., Chimbayo, A. & Najafi, N. (2003). A variable temperature, resonant density sensor made using an improved chip-level vacuum package. *Sensors and Actuators A* **107,** 119–124.
72. Svensson, K., Bülow, L., Kriz, D. & Krook, M. (2005). Investigation and evaluation of a method for determination of ethanol with the SIRE® Biosensor P100, using alcohol dehydrogenase as recognition element. *Biosensors and Bioelectronics* **21,** 705–711.
73. Taverniers, I., De Loose, M. & Van Bockstaele, E. (2004). Trends in quality in the analytical laboratory. II. Analytical method validation and quality assurance. *Trends in Analytical Chemistry* **23,** 535 – 552.
74. Urbano-Cuadrado, M., Luque de Castro, M.D., Pérez-Juan, P.M. & Gómez-Nieto, M.A. (2005). Comparison and joint use of near infrared spectroscopy and Fourier transform mid infrared spectroscopy for the determination of wine parameters. *Talanta,* **66**, 218-224.
75. Urbano-Cuadrado, M., Luque de Castro, M.D., Pérez-Juan, P.M., García-Olmo, J. & Gómez-Nieto, M.A. (2004). Near infrared reflectance spectroscopy and multivariate analysis in enology Determination or screening of fifteen parameters in different types of wines. *Analytica Chimica Acta* **527,** 81-88.
76. Urtubia, A., Pérez-Correa, J.R., Meurens, M. & Agosin, E. (2004). Monitoring large scale wine fermentations with infrared spectroscopy. *Talanta* **64,** 778–784.
77. Le Roux, J. (1997). Van Ryn advanced brandy course. The Van Ryn wine and spirit company.
78. Vlachos, N.A., Michail, C. & Sotiropoulou D. (2002). Is ISO/IEC 17025 Accreditation a Benefit or Hindrance to Testing Laboratories? The Greek Experience. *Journal of Food Composition and Analysis* **15,** 749–757.
79. Vonach, R., Lendl, B. & Kellner, R. (1998). High-performance liquid chromatography with real-time Fouriertransform infrared detection for the determination of carbohydrates, alcohols and organic acids in wines. *Journal of Chromatography A* **824**, 159-167.
80. Wold, S. (1995). Chemometrics; what do we mean with it, and what do we want from it? *Chemometrics and Intelligent Laboratory Systems* **30,** 109-115.
81. Yucesoy, D. & Ozen, B. (2013). Authentication of Turkish traditional aniseed flavoured distilled spirit, raki. *Food Chemistry* **141**, 1461 – 1465.
82. Zhang, Y.-L., Chen, J.-B., Lei, Y., Zhou, Q., Sun, S.-Q. & Noda, I. (2010). Discrimination of different red wine by Fourier-transform infrared and two-dimensional infrared correlation spectroscopy. *Journal of Molecular Structure* **974**, 144 – 150.
83. Zwinkels, J.C. (1996). Colour-measuring instruments and their calibration. *Displays* **16,** 163 – 171.

Chapter 3

Research results

Validation of method of analysis for ethanol, density, obscuration and colour of spirit products in a commercial laboratory in South Africa

CHAPTER 3 – Validation of method of analysis for ethanol, density, obscuration and colour of spirit products in a commercial laboratory in South Africa

ABSTRACT

Method validation is a necessary and important step to measure the precision, ruggedness, robustness, uncertainty and stability of analysis methods. This study validated the analytical methods for ethanol, density, obscuration and colour of South African spirit products. A distillation and oscillation-type density meter were evaluated for ethanol, density and obscuration analyses and a tintometer for colour analyses. Potstill and blended brandies were used on three different electric distilling points by one analyst to validate the ethanol and obscuration methods. Vodka, gin and brandies were used in the validation of the density method. One analyst analysed brandies with different gold intensities for the validation of the method of colour determination. Precision is expressed as the standard error of laboratory (SEL) and is compared with the precision reported for the commercial laboratory ($SEL_{Distell}$). The instruments were evaluated with the percentage relative standard deviation (%RSD) and analysis of variance (ANOVA) at 95% confidence level, as an indication of ruggedness and precision of the methods. The SEL of this study for brandy products using 100 mL sample volume is 0.097 %v/v and the %RSD of 1.65% indicates good precision. ANOVA tests confirmed 100 mL sample volumes gave statistical significantly more precise results than 50 mL sample volumes. The SEL were 0.00408 $^{20}_{20}$, 0.00312 $^{20}_{20}$ and 0.00031 $^{20}_{20}$, and the %RSD 0.14%, 0.10% and 0.01% for vodka, gin and brandy, respectively. Although the brandy results proved better precision, the ANOVA test proved no statistical significant difference between the analyses of the different products. Results for obscuration confirmed the more precise results using 100 mL samples. %RSD for obscuration with 100 mL sample volume is 1.65% and SEL 0.10. The repeatability of the colour analytical technique was found to be highly precise and repeatable. The SEL for colour is 0 gold units. The results obtained verify the use of these analytical techniques for future data collection.

3.1 INTRODUCTION

Method validation is the tool used to demonstrate that a specific analytical method measures what it is intended to measure, and is suitable for its intended purpose (ISO, 2005). It establishes the performance characteristics (Feinberg, 2007) to supply to clients and measures the different effects, throughout the whole analytical system, that influence the result (Taverniers *et al*., 2004) so analytical techniques can be compared. Accurate, precise and repeatable measurements are required throughout the production process. Products need to conform to internal specifications, measure up to the requirements of the consumers and comply with legal specifications (South African Liquor Products Act 60 of 1989).

The ethanol content of spirit products is a key analytical parameter in the distilled beverage industry. Ethanol concentration is highly regulated by governing bodies. In South Africa (SA) the Liquor Products Act (Act 60 of 1989) stipulates the legal ethanol concentration for each spirit product. Legal limitations prohibit bottling of spirit products below the legally allowed concentrations as stipulated in the SA Liquor Products Act (Act 60 of 1989). Taxes imposed on the volume of absolute ethanol keeps a tight rein on producers. The social and especially the economic implications of ethanol containing products require that the methods used for analyses are highly accurate, repeatable and precise.

The density of a material is defined as the mass per unit volume. Density plays an important role in final products for the accurate determination of volume. Producers are audited on legal requirements, stipulated in SANS 1841 (2008) in SA, on the average volume during bottling.

The total extract of spirits (in this case brandy) is normally expressed as the obscuration. Obscuration is of organoleptic importance as it is a measure of smoothness or sweetness, adding to the mouth-feel of a product.

The technique used in this study for ethanol, density and obscuration measurement is distillation with subsequent analysis with an oscillation-type density meter. Electronically controlled U-shaped oscillating metal or glass tubes and temperature control are used in the manufacturing of density meters. While the U-shaped tube is filled with a fluid, the tube is driven into resonance electrostatically and its motion sensed using metal electrodes placed under the microtube (Sparks *et al.,* 2003). The square of the resonance frequency is inversely proportional to the sum of the mass of tube and tube contents. As both the tube mass and tube inner volume are known values, the vibrating tube method allows the density of unknown fluids to be determined in a single measurement. The U-shaped tube is kept oscillating continuously at the characteristic frequency, which depends on the density of the filled-in sample. The oscillation period is measured and converted into density by the equation of the Mass-Spring-Model:

$$F = \frac{1}{2}\pi\sqrt{\frac{c}{(M+\rho V)}}$$

Where F is the frequency, c indicates the spring constant, M is the mass, ρ the density and V the volume (González-Rodríguez *et al*., 2003). The oscillation-type density meter is the method preferred by the alcoholic beverage industry (Patz *et al*., 2004; González-Rodríguez *et al.,* 2003; Lachenmeier *et al*., 2005; Lachenmeier, 2007; Lobo *et al*., 2006).

Colour is considered a major feature for the assessment of food product quality and is equally important for spirit products. Colour measurements are part of the final quality check of a product. Consistency in colour specific to each product must be maintained. A tintometer was used in this study for the colour measurement of spirit products. Tintometers have been used in studies on meat products (Sahoo and Anjaneyulu, 1997), vegetable oils (Subramanian *et al*., 1998; Hafidi *et al*., 2005), milkweed press oil (Holser, 2003), beer (Odibo *et al*., 2002) and on the bleaching power of bentonite on crude edible oils (Noyan *et al*., 2007).

The aim of this study was to validate the analytical methods of ethanol, density, obscuration and colour of SA spirit products in a routine commercial laboratory. Brandy, vodka and gin products are used in this validation study. Throughout this paper the standard error of laboratory of the commercial laboratory ($SEL_{Distell}$) will be compared with the results obtained in this study.

3.2 MATERIALS AND METHODS

3.2.1 Sample collection

Products prior to bottling were collected at a routine commercial laboratory of Distell (Distell Group (Pty) Ltd). A potstill and a blended brandy sample, as referred to in the SA Liquor Products Act 60 of 1989 (Act 60 of 1989), were used for the validation of ethanol and obscuration methods. The validation of the oscillation-type vibrating-U-tube density meter (model DMA-58, Anton Paar, Graz, Austria) was tested using gin, vodka and brandy samples. Four brandies with differing colour intensities were used for the colour validation. González and Herrador (2007) advise three to five replications for validations.

3.2.2 Methods validated

3.2.2.1 Ethanol concentration

The analysis method for ethanol concentration was according to the official analytical methods of the Association of Official Analytical Chemists (AOAC) (AOAC, 2005). Samples were distilled and subsequently analysed with an oscillation-type vibrating-U-tube density meter (model DMA-58, Anton Paar, Graz, Austria). Two different volumes of samples (50 mL and 100 mL) were evaluated for instrument precision when distilling the samples. Five replicates were done with two different sample volumes, two types of brandy and three electrical stills. These variables were included as a test of the ruggedness of the method. Figure 3.1 shows the experimental lay-out for the ethanol concentration method validation. Ethanol concentration results were reported to the third decimal.

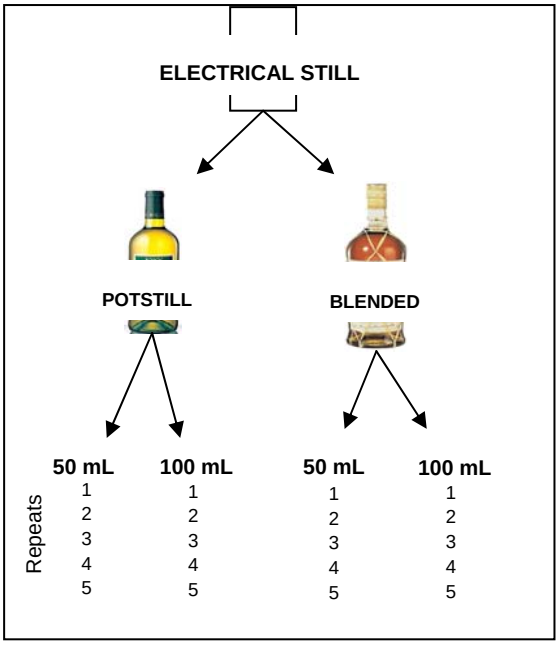


Figure 3.1: Experimental lay-out for ethanol and obscuration method validation done for each of the three electrical stills.

3.2.2.2 Density

An oscillation-type vibration-U-tube density meter (model DMA-58, Anton Paar, Graz, Austria) was used for the density method validation (AOAC, 2005). The precision for the density meter, reported as %RSD, was estimated by measuring repeatability on five replicate measurements with three different products (vodka, gin and brandy). This also gives an indication of the ruggedness by incorporating deliberate changes in the variables, in this case the three different products used. Density is highly dependent on temperature. The temperature (and unit) at which the analyses are taken is relative density of the sample at 20°C to density of water at 20°C. Density results were reported to the fifth decimal.

3.2.2.3 Obscuration

Obscuration of the samples was determined by calculating the difference between the true strength (TS) and direct strength (DS) of ethanol concentration of a sample. The TS is determined according to the protocol described under 3.2.2.1 and the ethanol concentration of the undistilled product determined by the density meter is the DS. The unitless obscuration results were reported to the second decimal.

3.2.2.4 Colour

A Lovibond Tintometer Model E was used for the colour measurements. The Lovibond tintometer used in this study arranges two adjacent fields of view, the product in the sample field and a white reflective surface in the comparison field. These are observed side by side. The analyst matches the colour of the sample with colour slides supplied with the tintometer. In this study, whisky and brandy samples are measured for colour in gold units. The precision and repeatability of the analyst were tested for the colour method, using four different brandies, with differing colour intensities. The analyst was presented with a set of five randomly ordered samples and asked to determine the colour of each sample. A set consisted of the four brandies and a repeat of one of the brandies. Four sets were given.

3.2.3 Statistical analysis

Dixon Q-tests were performed to identify any outliers in the data sets. Descriptive statistics in the Excel program (Microsoft Office 2000) was used to calculate performance parameters. The confidence interval presents the interval on the measurement scale within which the true value lies (Taverniers *et al.,* 2004). This study used a 95% probability.

The ruggedness of a method is evaluated by varying method parameters (Shabir, 2003) and measuring the degree of reproducibility under varying conditions (Taverniers *et al.,* 2004). In this study distilling point, sample volume in different products were the variables. Analysis of variance (ANOVA) tests were performed in the Excel program in the Microsoft Office 2000 package to measure the ruggedness of the method.

Measure of precision is given as the standard deviation (SD), the range and the % relative standard deviation (%RSD) (Mermet, 2008). The SD is the root mean square of deviation from the mean of the set with *n* number of samples calculated with the equation:

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

where x_i is item *i* in the set, $\bar{x}$ is the mean of the set and *n* is the number of samples.

%RSD is determined with the equation:

$$\%RSD = \left(\frac{SD}{Mean} \right) \times 100$$

Where the *mean* is calculated as the sum of the variable values, divided by the number of samples (Taverniers *et al*., 2004). The %RSD is used as a standard procedure to measure instrument precision and has been used in various wine quantification FT-IR studies (Soriano *et al*., 2007; La Torre *et al*., 2006; Lachenmeier *et al.,* 2007). Acceptable %RSD for analytes with units of 1% (ethanol, obscuration and colour) is 2.7% and analytes with units of 0.1% is 3.7% (density) (González and Herrador, 2007).

The precision of the methods is expressed as the Standard Error of Laboratory (SEL). SEL was calculated as in the following equation:

$$SEL = \sqrt{\frac{\Sigma (y_1 - y_2)^2}{2n}}$$

Where y_1 and y_2 are the values from duplicate determinations and n is the number of samples. The reported $SEL_{Distell}$ is the precision of Distell laboratories (Work instruction of in-house validation of results, Distell Group (Pty) Ltd, SA) compared to the precision proven in this study.

Since this study is performed in a controlled air-conditioned laboratory the validation of the robustness of the methods were not necessary as these conditions did not affect the methods of analyses. All the statistical results were reported to the second decimal.

3.3 RESULTS AND DISCUSSION

3.3.1 Ethanol analysis method validation

The results and statistical descriptors for the validation of the ethanol concentration method using potstill brandy and blended brandy are given in Table 3.1. No outliers were detected. The %RSDs over all the distilling points indicate good precision (González and Herrador, 2007) ranging from 0.08 to 0.33% for potstill brandy, and 0.05 to 0.15% for blended brandy. The confidence interval gives the interval within which the true value lies with a 95% probability.

It was noted that the 100 mL sample volume gave higher ethanol concentrations at every distilling point with both brandy types. The difference in the ethanol concentration results between the 50 mL and 100 mL brandy samples can be explained by scrutinising the ratio of the actual amount of water and ethanol in the round bottom flasks with distillation compared to the amount of distillate obtained. With both volumes distilled, the volume of distillate captured after cooling was not in ratio to the starting volume. The resulting effect is the distillate fraction in the 100 mL brandy sample is less compared to the starting volume, and therefore in effect a more volatile (higher in ethanol) fraction. Alternatively when the 50 mL samples were distilled, the distillate contained a higher percentage higher boiling compounds (water), effectively thinning the ethanol concentration in the sample.

The average SEL over all the distilling points for 100 mL sample volume using both types of brandy is 0.097 %v/v. The $SEL_{Distell}$ for ethanol concentration determination for spirit products is 0.200 %v/v. The decrease in precision with the Distell error is due to the inclusion of variables such as different analysts, different spirit products, different laboratories, different distilling equipment, etc.

One-way ANOVA with a 95% confidence level was performed to validate the ruggedness of the method. The null hypothesis tested was that there are no statistical differences between the distilling points. The results are given in Table 3.2. If the $F_{calculated}$ value exceeds the $F_{critical}$ value,

the null hypothesis is rejected at the 95% level of significance. Statistical significant differences ($F_{calculated} > F_{critical}$) were found between the distilling points with 50 mL sample volume of the blended brandy and 100 mL sample volume of the potstill brandy, therefore the method is not rugged over the distilling points and the results obtained with the replicates verify the precision of the execution of the reference method.

Table 3.1: Ethanol concentrations (%v/v) obtained with distillation and subsequent density meter analysis of two different sample volumes of potstill brandy and blended brandy using three different electric distilling points. Statistical descriptors for the results are also given.

			Distilling point					
			1		*2*		*3*	
			50 mL	100 mL	50 mL	100 mL	50 mL	100 mL
POTSTILL BRANDY	Replicates	1	38.032	38.225	38.139	38.141	38.079	38.216
		2	38.074	38.180	38.155	38.153	38.067	38.259
		3	37.984	38.213	37.884	38.146	38.059	38.287
		4	38.050	38.238	38.200	38.173	38.111	38.350
		5	38.059	38.146	38.131	38.216	38.160	38.243
	Descriptive Statistics	Range	37.984–38.074	38.146–38.238	37.884–38.200	38.141–38.216	38.059–38.160	38.216–38.350
		SD[a]	0.04	0.04	0.13	0.03	0.04	0.05
		%RSD[b]	0.09	0.10	0.33	0.08	0.11	0.13
		SEL[c]	0.110	0.118	0.394	0.097	0.131	0.162
		95% CI[d]	38.04±0.04	38.200±0.05	38.102±0.16	38.166±0.04	38.095±0.05	38.271±0.06
BLENDED BRANDY	Replicates	1	41.212	41.470	41.413	41.474	41.357	41.497
		2	41.355	41.517	41.413	41.507	41.477	41.543
		3	41.295	41.505	41.507	41.504	41.380	41.527
		4	41.289	41.470	41.352	41.533	41.460	41.539
		5	41.300	41.481	41.358	41.529	41.423	41.513
	Descriptive Statistics	Range	41.212–41.355	41.470–41.517	41.352–41.507	41.474–41.533	41.357–41.477	41.497–41.543
		SD[a]	0.05	0.02	0.06	0.02	0.05	0.02
		%RSD[b]	0.12	0.05	0.15	0.06	0.12	0.05
		SEL[c]	0.162	0.068	0.197	0.075	0.161	0.060
		95% CI[d]	41.290±0.06	41.489±0.03	41.409±0.08	41.509±0.03	41.419±0.06	41.524±0.02

[a]SD - standard deviation; [b]%RSD - % relative standard deviation; [c]SEL - standard error of laboratory; [d]95% CI - confidence interval at 95%

Table 3.2: Summary of one-way ANOVA tests to examine differences if any, between stills for the potstill brandy and blended brandy.

	Source of Variation	*SS*[a]	*df*[b]	*MS*[c]	$F_{calculated}$	*P-value*	$F_{critical}$
50 mL sample volume	Between distilling points for potstill brandy	0.01	2	0.01	0.94	0.42	3.885
	Between distilling points for blended brandy	0.05	2	0.03	8.48	0.005	3.885
100 mL sample volume	Between distilling points for potstill brandy	0.03	2	0.01	8.74	0.005	3.885
	Between distilling points for blended brandy	0.003	2	0.002	3.42	0.07	3.885

[a]SS – sum of squares; [b]df – degrees of freedom; [c]MS – mean square

A two-way ANOVA was done to investigate possible interaction between sample volume and distilling point (Table 3.3). The F-test demonstrated the difference between the sample volumes previously noted are statistical significant ($F_{calculated} > F_{critical}$) for both potstill and blended brandy. For the potstill brandy the two-way ANOVA showed no statistical significant difference between the distilling point and no interaction between distilling point and sample volume. For the blended brandy statistical significant difference was found between distilling points and the interaction between the distilling point and sample volume was significant. From the ANOVA results the ruggedness of the method could not be validated.

Table 3.3: Two-way ANOVA to examine the interaction between the distilling point and sample volume for the potstill and blended brandy.

	Source of Variation	*SS*[a]	*df*[b]	*MS*[c]	$F_{calculated}$	*P-value*	$F_{critical}$
Potstill	Between 50 mL and 100 mL sample volume	0.13	1	0.13	34.28	4.87E-06	4.260
	Between distilling points	0.02	2	0.01	2.82	0.08	3.403
	Interaction between distilling point and sample volume	0.02	2	0.01	2.36	0.12	3.403
Blended	Between 50 mL and 100 mL sample volume	0.14	1	0.14	77.86	5.34E-09	4.260
	Between distilling points	0.04	2	0.02	11.24	3.59E-04	3.403
	Interaction between distilling point and sample volume	0.015	2	0.008	4.391	0.024	3.403

[a]SS – sum of squares; [b]df – degrees of freedom; [c]MS – mean square

Although there was different outcomes for the two different brandies regarding the precision of the two sample volumes, the average error for the 100 mL sample volume was smaller than the average error over all the 50 mL sample volumes. Subsequently 100 mL sample volume will be used in further analyses as this gave statistical significant more precise results. The ruggedness could not be validated between distilling points. The %RSDs over all the distilling points, <2.7%, indicate good precision (González and Herrador, 2007). The SEL is 0.097 %v/v and $SEL_{Distell}$ is 0.200 %.

3.3.2 Density analysis method validation

The results and statistical descriptors for the validation of the density analysis method using vodka, gin and brandy are given in Table 3.4. No outliers were detected. The %RSD values for the density measurements illustrate good precision (<3.7%) and reproducibility of the analytical method (González and Herrador, 2007).

Table 3.4: Density values ($^{20}_{20}$) obtained with density meter analyses of three different spirit products.

Replicates	Vodka	Gin	Brandy
1	0.94297	0.94311	0.94476
2	0.94469	0.94302	0.94470
3	0.94469	0.94482	0.94492
4	0.94502	0.94497	0.94492
5	0.94660	0.94479	0.94480
Range	*0.94297 – 0.94660*	*0.94302 – 0.94497*	*0.94470 – 0.94492*
SD[a]	*0.00129*	*0.00099*	*0.00010*
%RSD[b]	*0.14*	*0.10*	*0.01*
SEL[c]	*0.00408*	*0.00312*	*0.00031*
95% CI[d]	*0.94479±0.002*	*0.94414±0.001*	*0.94482±0.0001*

[a]SD - standard deviation; [b]%RSD - % relative standard deviation; [c]SEL - standard error of laboratory; [d]95% CI - confidence interval at 95%

One-way ANOVA test was done in order to evaluate the ruggedness of the method. The null hypothesis tested was: there are no statistical significant differences between the ability of the method to measure the different spirit products. The $F_{calculated}$ value of 0.84 was smaller than the $F_{critical}$ value at $P=0.05_{2;12}$ of 3.885; accepting the null hypothesis. Although the brandy showed better repeatability than the clear spirits, the difference is not statistical significant.

The SEL of the brandy product (0.00031 $^{20}_{20}$) is the only error conforming to the $SEL_{Distell}$ of 0.001 $^{20}_{20}$. When using density measurements for ethanol concentration analysis, the error is multiplied. A 0.0001 g/mL difference in density equals to a difference of 0.07 %v/v in the ethanol concentration. A difference of 0.01 g/mL equals to a difference of 7.19 %v/v! Although the clear spirit products showed lower precision, the ANOVA test proved no statistical significant difference of analyses between the products; thus this method is proven suitable for all the products. The confidence interval gives the interval within which the true value lies with a 95% probability.

3.3.3 Obscuration analysis method validation

The DS for potstill brandy was determined as 35.781 %v/v and 40.235 %v/v for the blended brandy. The TS values are given in Table 3.1; the resultant obscuration values are given in Table 3.5.

No outliers were detected. The %RSD for the obscuration determination of the potstill brandy with 50 mL sample volume using distilling point 2 as well as all three distilling points with the blended brandy fell outside the 2.7% benchmark (González and Herrador, 2007) and indicated poor precision. This supports the decision to use 100 mL sample volume for quantification from the ethanol concentration validation results. The $SEL_{Distell}$ for the quantification of obscuration is 0.2 while this validation indicates a SEL = 0.10. As in the ethanol validation results, there is an increase in the precision validated in this study compared to the Distell internal validation, accounted to more variability included in the Distell internal validation. The confidence interval gives the interval within which the true value lies with a 95% probability.

Table 3.5: Obscuration values obtained by the difference between the TS and DS of two different sample volumes of potstill brandy and blended brandy using three different electric distilling points.

			Distilling point					
			1		*2*		*3*	
			50 mL	100 mL	50 mL	100 mL	50 mL	100 mL
POTSTILL BRANDY DS = 35.781 %V/V	Replicates	1	2.25	2.44	2.36	2.36	2.30	2.44
		2	2.29	2.40	2.37	2.37	2.29	2.48
		3	2.20	2.43	2.10	2.37	2.28	2.51
		4	2.27	2.46	2.42	2.39	2.33	2.57
		5	2.28	2.37	2.35	2.44	2.38	2.46
	Descriptive Statistics	Range	2.20 – 2.29	2.37 – 2.46	2.10 – 2.42	2.36 – 2.44	2.28 – 2.38	2.44 – 2.57
		SD[a]	0.04	0.04	0.13	0.03	0.04	0.05
		%RSD[b]	1.54	1.54	5.37	1.28	1.78	2.05
		SEL[c]	0.11	0.12	0.39	0.10	0.13	0.16
		95% CI[d]	2.26±0.04	2.42±0.05	2.32±0.16	2.39±0.04	2.31±0.05	2.49±0.06
BLENDED BRANDY DS = 40.235 %V/V	Replicates	1	0.98	1.24	1.18	1.24	1.12	1.26
		2	1.12	1.28	1.18	1.27	1.24	1.31
		3	1.06	1.27	1.27	1.27	1.15	1.29
		4	1.05	1.24	1.12	1.30	1.23	1.30
		5	1.07	1.25	1.12	1.29	1.19	1.28
	Descriptive Statistics	Range	0.98 – 1.12	1.24 – 1.28	1.12 – 1.27	1.24 – 1.30	1.12 – 1.24	1.26 – 1.31
		SD[a]	0.05	0.02	0.06	0.02	0.05	0.02
		%RSD[b]	4.84	1.70	5.30	1.85	4.31	1.48
		SEL[c]	0.16	0.07	0.20	0.08	0.16	0.06
		95% CI[d]	1.06±0.06	1.25±0.03	1.17±0.08	1.27±0.03	1.18±0.06	1.29±0.02

[a]SD - standard deviation; [b]%RSD - % relative standard deviation; [c]SEL - standard error of laboratory; [d]95% CI - confidence interval at 95%

3.3.4 Colour analysis method validation

Results for the validation for the test for colour in spirit products are given in Table 3.6.

Table 3.6: Colour values (gold units) obtained with a tintometer of four different brandies by one analyst.

Replicates	Sample A	Sample B	Sample C	Sample D
1	20	13	11	9
2	20	13	11	9
3	20	13	11	9
4	20	13	11	9
5	20	13	11	9
Range	20 - 20	13 - 13	11 - 11	9 - 9
SD[a]	0	0	0	0
%RSD[b]	0	0	0	0
SEL[c]	0	0	0	0
95% CI[d]	20 ± 0	13 ± 0	11 ± 0	9 ± 0

[a]SD - standard deviation; [b]%RSD - % relative standard deviation; [c]SEL - standard error of laboratory; [d]95% CI - confidence interval at 95%

The results prove this method of analysis is highly precise, accurate and repeatable with an error of 0 gold units. The shortcoming of this method is it is highly subjective, relying on the analyst's interpretation and experience. The $SEL_{Distell}$ is 2 gold units, which includes variables such as different analysts, different laboratories, different products, etc. The confidence interval gives the interval within which the true value lies with a 95% probability.

3.4 CONCLUSION

Method validation was performed to test the precision, ruggedness and robustness for a specific product over the specified range and provides an assurance of reliability. The ethanol validation demonstrated that the 100 mL sample volume was more precise and gave better repeatability throughout the validation. Two-way ANOVA tests proved that 100 mL sample volume gave statistical significantly more precise results than 50 mL sample volume. The outcome is that only 100 mL sample volumes will be used in subsequent analyses for ethanol and obscuration measurements. The %RSD values of the methods indicate good precision for ethanol with 100 mL sample volume, averaged over distilling points and the two types of brandy (%RSD = 0.08); density with %RSD of 0.14%, 0.10% and 0.01% for vodka, gin and brandy respectively; obscuration with 100 mL sample volume, averaged over distilling points and the two types of brandy (%RSD = 1.65%) and colour (%RSD = 0%). ANOVA tests could not validate the ruggedness of the method of analysis for ethanol concentration. Statistical significant differences were found between sample volume and distilling points. The precision of the methods were validated and indicated with the SEL and compared to the $SEL_{Distell}$. The ethanol (SEL = 0.097 %v/v and $SEL_{Distell}$.= 0.200 %v/v), obscuration (SEL = 0.10 and $SEL_{Distell}$.= 0.2) and colour (SEL = 0 gold units and $SEL_{Distell}$.= 2 gold units) showed an improvement on the $SEL_{Distell}$. The increase in precision is due to the exclusion of variables included in the validation study by the Distell laboratory. In the case of the density, the SEL for the brandy product 0.00031 $^{20}_{20}$ was an improvement on the $SEL_{Distell}$ of 0.001 $^{20}_{20}$. Although the vodka and gin products showed a decrease in precision according the SEL (0.00408 $^{20}_{20}$ and 0.00312 $^{20}_{20}$ respectively) compared with the $SEL_{Distell}$, the ANOVA proved no statistical significant difference between the results of the three products. The subjective colour analysis method gave impeccable results for precision. Comparing this validation results to the Distell SEL, a decrease in SEL was expected as the Distell SEL was calculated over various laboratories, including another variable into the equation.

3.5 REFERENCES

1. Association of Official Analytical Chemists (AOAC). (2005). Official methods of analysis of the Association of Official Analytical Chemists, *Association of Official Analytical Chemists*, Arlington, VA.
2. Feinberg, M. (2007). Validation of analytical methods based on accuracy profiles. *Journal of Chromatography A* **1158,** 174-183.
3. González, A.G. & Herrador, M.A. (2007). A practical guide to analytical method validation, including measurement uncertainty and accuracy profiles. *Trends in Analytical Chemistry* **26,** 227-238.
4. González-Rodríguez, J., Pérez-Juan, P. & Luque de Castro, M.D. (2003). Determination of ethanol in beverages by flow injection, pervaporation and density measurements. *Talanta* **59**, 691-696.
5. Hafidi, A., Pioch, D. & Ajana, H. (2005). Membrane-based simultaneous degumming and deacidification of vegetable oils. *Innovative Food Science and Emerging Technologies* **6**, 203-212.
6. Holser, R.A. (2003). Properties of refined milkweed press oil. *Industrial Crops and Products* **18**, 133-138.
7. International Organization for Standardization (ISO), ISO/IEC 17025. (2005). General requirements for the competence of testing and calibration laboratories, ISO, Geneva, Switserland.
8. La Torre, G.L., Saitta, M., Vilasi, F., Pellicano, T. & Dugo, G. (2006). Direct determination of phenolic compounds in Sicilian wines by liquid chromatography with PDA and MS detection. *Food Chemistry* **94**, 640-650.
9. Lachenmeier, D.W. (2007). Rapid quality control of spirit drinks and beer using multivariate data analysis of Fourier transform infrared spectra. *Food Chemistry* **101**, 825-832.
10. Lachenmeier, D.W., Burri, P.A., Fauser, T., Frank, W. & Walch, S.G. (2005). Rapid determination of alcoholic strength of egg liquer using steam distillation and oscillation-type densimetry with peristaltic pumping. *Analytica Chimica Acta* **537**, 377-384.
11. Lobo, A.P., Valles, B.S., Tascón, N.F., Madrera R.R. & García, O.F. (2006). Calibration models for routine analysis of cider by mid-infrared spectroscopy. *LWT-Food Science & Technology* **39,** 1026-1032.
12. Mermet, J.-M. (2008). Limit of quantitation in atomic spectrometry: An unambiguous concept? *Spectrochimica Acta Part B* **63,** 166-182.
13. Noyan, H., Önal, M. & Sarikaya, Y. (2007). The effect of sulphuric acid activation on the crystallinity, surface area, porosity, surface acidity, and bleaching power of a bentonite. *Food Chemistry* **105**, 156-163.
14. Odibo, F.J.C., Nwankwo, L.N. & Agu, R.C. (2002). Production of malt extract and beer from Nigerian sorghum varieties. *Process Biochemistry* **37**, 851-855.
15. Patz, C.-D., Blieke, A., Ristow, R. & Dietrich, H. (2004). Application of FT-MIR spectrometry in wine analysis. *Analytica Chimica Acta* **513**, 81-89.
16. Sahoo, J. & Anjaneyulu, A.S.R. (1997). Quality Improvement of Ground Buffalo Meat by Preblending with Sodium Ascorbate. *Meat Science* **46**, 237-247.
17. Shabir, G.A. (2003). Validation of high-performance liquid chromatography methods for pharmaceutical analysis. Understanding the differences and similarities between validation requirements of the US Food and Drug Administration, the US Pharmacopeia and the International Conference on Harmonization. *Journal of Chromatography A* **987**, 57–66.
18. Soriano, A., Pèrez-Juan, P.M., Vicario, A., González, J.M. & Pèrez-Coello, M.S. (2007). Determination of anthocyanins in red wine using newly developed method based on Fourier transform infrared spectroscopy. *Food Chemistry* **104**, 1295-1303.

19. South African National Standard. SANS1841:2008. Control of the quantity of contents in prepacked packages within the prescriptions of legal metrology legislation.
20. Sparks, D., Smith, R., Schneider, R., Cripe, J., Massoud-Ansari, S., Chimbayo, A. & Najafi, N. (2003). A variable temperature, resonant density sensor made using an improved chip-level vacuum package. *Sensors and Actuators A* **107,** 119–124.
21. Subramanian, R., Nakajima, M., Kimura, T. & Maekawa, T. (1998). Membrane process for premium quality expeller-pressed vegetable oils. *Food Research International* **31**, 587-593.
22. Taverniers, I., De Loose, M. & Van Bockstaele, E. (2004). Trends in quality in the analytical laboratory. II. Analytical method validation and quality assurance. *Trends in Analytical Chemistry* **23**, 535 – 552.
23. Taverniers, I., Van Bockstaele, E. & De Loose, M. (2004). Trends in quality in the analytical laboratory. I. Traceability and measurement uncertainty of analytical results. *Trends in Analytical Chemistry* **23**, 480-490.
24. South African Liquor Products Act No. 60 of 1989.

Chapter 4

RESEARCH RESULTS

Quantification of ethanol, density, obscuration and colour of spirit products using Fourier transform infrared (FT-IR) spectroscopy and multivariate data analysis methods

CHAPTER 4 – Quantification of ethanol, density, obscuration and colour of spirit products using Fourier transform infrared (FT-IR) spectroscopy and multivariate data analysis methods

ABSTRACT

Fourier transform infrared (FT-IR) spectroscopy coupled with multivariate data analysis methods is a proven powerful analytical tool in the wine industry. However, the technology has seen limited quantitative application for spirit products in South Africa. The objective of this study was, therefore, to establish calibration models for the quantification of ethanol, density, obscuration and colour in South African spirit products, using FT-IR spectroscopy in combination with multivariate data analysis**.** A sample set consisting of 363 samples and the following product types, brandies (273), whiskies (70) gins (14) and vodkas (6), was collected. Infrared spectra of the samples in the range 926 to 5012 cm^{-1} were collected and reference data for ethanol, density, obscuration and colour were obtained with appropriate validated reference methods. Principal component analysis (PCA) was applied to the data to explore the data structure and to identify outliers. Partial least squares (PLS) regression was applied to construct calibration models for ethanol, density, obscuration and colour. Independent test set validation was used to evaluate the performance of the respective models. Root mean square error of prediction (RMSEP) values for the respective calibration models were in agreement with the standard errors of the laboratory methods. The RMSEP values for the ethanol calibrations models with the wavenumber selection per product (RMSEP = 0.038 to 0.106 %v/v), exceeded and showed significant improvement on previously reported prediction errors. The results for the density calibration showed a similar trend, with the product specific calibration models outperforming the calibration model when all samples were used in one set. This study produced novel results for quantification of obscuration (RMSEP = 0.10 and 0.09 in blended brandies and potstill brandies, respectively) and colour (RMSEP < 2.286 gold units) of brandies and whiskies. The correlation coefficients (R^2) between true and predicted values, for the four parameters tested, indicated good to excellent precision ($0.8<R^2<1.0$). The results confirmed FT-IR spectroscopy as a method for the quantification and/or screening of ethanol, density, obscuration and colour of spirit products. This report is the first for spirit products in the South African wine industry and a pioneer study on the quantification of obscuration and colour of spirit products with FT-IR spectroscopy and multivariate data analysis. The impact of the results obtained in this study is that the respective PLS models that were developed can now be transferred to industry for quality control checks of large volumes of spirits products that are routinely sent to the analysing laboratories.

4.1 INTRODUCTION

Quality control of spirit products involves the review, maintenance and insurance of the product adherence to specifications previously defined. This incorporates different analyses at different stages of the production process. Difficulty, objectivity and time to complete the analyses were the base for choosing ethanol, obscuration, density and colour for this study. Distillation of the sample with subsequent analysis with an oscillation type density meter was the approved reference method for a Distell laboratory for ethanol and obscuration analyses. Two steps to the method of analysis for ethanol and obscuration allows for a greater opportunity for experimental error. One obscuration measurement can take up to 40 minutes to complete. A tintometer (using colourimetry) was the reference method used for the routine analysis of colour in spirit products. This is an instrument that relies on the human eye as the detector, comparing the sample with colour standards (Skoog *et al.,* 1998). These results can be subjective and biased to personal experience and preferences of the analyst. Therefore, for a standard analysis of these four parameters for one spirit sample, four different measuring techniques are used, that is time consuming and expensive. There is a need in the South African spirit product industry for a technique that is fast, accurate and reliable.

Fourier transform infrared (FT-IR) spectroscopy is a proven tool in wine industries worldwide for quantification purposes. Nieuwoudt *et al.* (2006) evaluated fermentation profiles of *Saccharomyces cerevisiae* in fermenting Chenin blanc musts and excellent quantitative results for various parameters of interest, were obtained. Several studies reported on the determination of anthocyanins in red wine, using FT-IR spectroscopy (Soriano *et al.,* 2007; Romera-Fernández *et al.,* 2012). Despite widespread implementation in wine chemical profiling (Patz *et al.,* 2004; Cozzolino*,* 2012), the application of FT-IR spectroscopy to spirit products in South Africa remains minimal.

Developments in infrared spectroscopy instrumentation and the availability of chemometric software on the instrumentation that can correlate spectra with chemical compositional data, offer a one-step solution, which gives multi-component tests in one single analysis, all in a fraction of the time required by current traditional chemical methods of analysis. However, to the best of our knowledge the application of FT-IR spectroscopic analyses to South African spirit products have not been implemented on the scale required, mainly as a result of the absence of validated PLS models. Furthermore, no studies have been reported for analysis of the colour of spirit products, which is an important quality indicator in brandy and whisky where it contributes to the organoleptic characterisation of products. Previous studies on ethanol quantification in spirit products with FT-IR spectroscopy reported the requirement of prior dilution of the samples (Gallignani *et al.,* 1994, 2005), thereby increasing the analysis time significantly. Lachenmeier (2007) proved FT-IR spectroscopy as a fast and reliable means of spirit product analyses. Density, ethanol, methanol, ethyl acetate, propanol-1, isobutanol and 2-/3-methyl-1-butanol were quantified with great accuracy with PLS calibration models (Lachenmeier, 2007).

The objective of this project was to build PLS calibration models for the quantification of ethanol, density, obscuration and colour for South African spirit products with FT-IR spectroscopy, using instrumentation that is widely employed by analytical laboratories in South African wine laboratories.

4.2 MATERIALS AND METHODS

4.2.1 Samples

A total of 363 samples were collected at Distell (Distell Group (Pty) Ltd) routine laboratories during 2005, 2006 and 2009. The set included 273 brandy samples, 70 whisky samples, 14 gin samples and 6 vodka samples. Twenty two different brands were included in this study. Due to the commercial nature of this study, the samples submitted to the routine laboratory were the samples used in this study.

4.2.2 Reference methods

All reference methods used were validated for the purpose of this study as described in Chapter 3. Reference values for ethanol and density were obtained according to the official analytical methods of the Association of Official Analytical Chemists (AOAC, 2005). Ethanol concentration (%v/v) was determined by distilling 100 mL aliquots of each sample. The distillate was collected for subsequent quantification of the ethanol concentration with an oscillation-type vibrating-U-tube density meter (model DMA-58, Anton Paar, Graz, Austria). The unit for density measurements in this study ($^{20}_{20}$) is relative density of the sample at 20°C compared to density of water at 20°C. Obscuration (a unit less measurement) was determined by calculating the difference between the true strength (TS) and direct strength (DS) of ethanol concentration of a sample. The TS ethanol concentration was determined by distilling 100 mL of the sample and subsequent quantification with an oscillation-type density meter. The DS ethanol concentration was the ethanol concentration of the undistilled product determined by the density meter. Obscuration values were only obtained for brandy products. Other products in this study did not have a value for obscuration, or the values were so small, that they were not taken into account. Colour measurements were done using a Lovibond Tintometer Model E (Camlab, UK). Only whisky and brandy products were analysed for colour measurement that were reported in gold units. The other products used in this study, gin and vodka, are colourless.

4.2.3 FT-IR spectral acquisition

The spectral data were collected with a WineScan FT120 instrument (Foss Analytical, Denmark). The instrument was equipped with an autosampler (64 tray, 30 mL cups), model 5027 (Foss Analytical, Denmark). No prior sample preparation was done before spectral acquisition. A sample volume of 30 mL was pumped through an in-line CaF_2-lined cuvette, with optical path length of 37 µm, which was kept at 40°C. Spectra were generated from 929.78 to 4996.11 cm^{-1}. Duplicate scans of the same sample were obtained. The apparatus was cleaned with the S-470 Cleaning Agent Solution (Foss Analytical, Denmark, 2001), and zeroed automatically every 120 minutes with the Zero solution (Foss Analytical, Denmark, 2001). Duplicate spectra were exported to the Unscrambler Software (version 9.2, Camo ASA, Trondheim, Norway) averaged at each wavenumber to give a single spectrum per sample and averaged spectra were exported to the SIMCA Software (version 13.0.3, Umetrics, Malmö, Sweden). The Unscrambler software was used to construct the calibration models while the SIMCA software was used in the initial explorative data analysis. Visual inspection of the spectra was done to identify any samples that showed uncharacteristic properties.

4.2.4 Data analysis

4.2.4.1 Explorative data analysis

Principal component analysis (PCA) was performed prior to calibration model development, to investigate similarities, differences, groupings and patterns between the observations (Esbensen, 2002). With PCA, the dimensionality of the data is reduced to a small number of principal components (PC's), which are linear functions of the original variables (Esbensen, 2002). The score plot shows the location of the observations along each component. In conjunction with the score plots, the loading plots give information on the importance of the variables in relation to the variation in the data seen in the score plot. Water absorbing regions in infrared spectra are associated with noise and these were deselected for analysis. The spectral data was mean-centred, so that all the components have the same centroid as their origin, to ensure a tight model (Craig *et al.,* 2006). The **y**-variables were unit variance (UV) scaled, which involves multiplying the reference **y**-variable value with 1/SD (Craig *et al.,* 2006). Hotelling T^2 ellipse at a 95% confidence level on the scores plots, display a measure to indicate how far each observation is from the center of the model (Esbensen, 2002).Observations far outside the ellipse can be considered as serious outliers and should be investigated further.

4.2.4.2 PLS regression modelling

Partial least squares (PLS) regression is a modelling technique that establishes a relationship between **x**- and **y**-variables (Esbensen, 2002). In this study the **x**-variables were infrared spectra and the **y**-variables were ethanol, density, obscuration and colour.

Two options were tested for the calibration model sample set selection. Firstly, all the different spirit products were included into one calibration and secondly, calibration models according to product type were developed to minimise the variation between the calibration samples with the possibility of minimising the prediction error. All the products in one calibration could lead to higher prediction errors, due to the high variability between the samples, but could result in a more robust model that can be used for all the spirit products. The latter option is attractive from the laboratory perspective, since the scanning procedure is simplified by eliminating the decision-step about which model to use, prior to analysis.

Additionally, with regards to the spectra used in the calibration models, three strategies were investigated. *Strategy 1* was the raw spectra with water regions deselected. *Strategy 2* was spectral pre-processing by means of Savitzky-Golay (S.G.) transformation and smoothing (five smoothing points and second polynomial order filtering). S.G. transformation has been used in various classification and quantification studies in the alcoholic beverage industry (Pontes *et al.,* 2006; Cozzolino *et al.,* 2007; Di Egidio *et al.,* 2011; McIntyre *et al.,* 2011). The water regions and uninformative wavenumbers were removed after derivatisation. *Strategy 3* was a wavenumber selection strategy on the pre-processed spectra, which was based on the assumption that the Beer-Lambert law (Silverstein *et al.,* 1991) was true for this data set, which has been used in other quantification studies on wine (Moreira and Santos, 2004).

Only test set validation was done and not resubstitution to evaluate the performance of all the calibration models developed in this study. One third of the samples per data set were randomly chosen as a validation set, to assess the prediction accuracies of the models. Sample outliers were identified with X-Y relation plots for each model. This is a plot of t-scores versus u-scores. Samples that lie off the regression line are possible outliers. Outliers were removed from further analyses.

4.2.4.3 Statistics used to evaluate the calibration models

Statistics calculated for the calibration models were correlation coefficients (R^2) (see Table 4.1 for interpretations), root mean squared error of cross validation (RMSECV), root mean squared error of prediction (RMSEP) and residual predictive deviation (RPD) (see Table 4.1 for interpretations). R^2 is an evaluation measure of the precision of the calibration model. RPD, defined as the ratio between the standard deviation (SD) of the reference data and the RMSEP for the predictions (RPD = SD/RMSEP) (Pink *et al.,* 1998), was applied to evaluate the prediction ability of a calibration model.

RMSECV is obtained by the equation:

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

where $\hat{y}_i$ denotes predicted concentration, y_i the actual concentration of the *i*th sample, as determined by the reference method, and *n* the number of samples used in the calibration model (Esbensen, 2002). When the prediction performance of the validation set is assessed, the equation is replaced by RMSEP, where *n* is the number of samples used in the validation set. The RMSEP value gives an indication of the expected error for future predictions and is compared to the standard error of laboratory (SEL). The Distell inter-laboratory SEL ($SEL_{Distell}$) as well as own (this researcher) validation data ($SEL_{own\ data}$) (see Chapter 3) were used in the comparison. The two SEL values were compared in order to evaluate the error added by different analysts and laboratories. In this study the $SEL_{own\ data}$ was one analyst in one laboratory, compared to the $SEL_{Distell}$ which was measured over time and pertained to various analysts at various laboratories with various product types. The RMSEP value was compared with the SEL values to compare the average error of the predicted values with that obtained with the reference analysis methods.

Table 4.1: Criteria for the evaluation of R^2 (adapted from Urbano-Cuadrado *et al*., 2005) and RPD (Pink *et al.,* 1998) used to evaluate the performance of the calibration models.

R^2 [a]	**Interpretation**
≥ 0.90	Excellent precision
0.70 – 0.89	Good precision
0.50 – 0.69	Good separation between low, medium, and high values
0.30 – 0.49	Correct separation between low and high values
0.05 – 0.29	Not suitable for quantification
RPD [b]	**Interpretation**
<3	Calibration model unsuitable for accurate quantification
3 – 5	Calibration model suitable for screening purposes
>5	Calibration model suitable for quantification

[a]R^2 - correlation coefficient; [b]RPD - residual predictive deviation

4.3 RESULTS AND DISCUSSION

4.3.1 Overview of the data set

Descriptive statistics of the sample sets per parameter are summarised in Table 4.2. Ethanol measurements were determined to an accuracy of three decimal places, while the density measurements were determined to an accuracy of five decimal places. Density and ethanol are

inversely correlated with each other. Obscuration is determined to an accuracy of two decimals, as well as the colour measurements, measured in gold units only, since the coloured spirit products used in this study, were gold toned.

Table 4.2: Descriptive statistics of the samples used for building PLS calibration models for ethanol, density, obscuration and colour of spirit products.

Parameter	N[a]	SD[b]	Skew-ness	$SEL_{Distell}$[c]	$SEL_{own\ data}$[d]	Product types	Minimum	Maximum	Mean
Ethanol (%v/v)	363	1.95	-2.14	0.200	0.097	PB[e] (44)	33.748	43.556	38.008
						BB[f] (229)	38.060	46.000	42.987
						WH[g] (70)	39.952	44.730	43.099
						VO[h] (6)	42.950	43.320	43.153
						GI[i] (14)	35.390	43.390	42.674
Density ($^{20}_{20}$)	187	0.00296	1.98	0.001	$0.00360_{clear\ spirits}$ $0.00031_{brandies}$	PB[e] (16)	0.94600	0.95472	0.95344
						BB[f] (141)	0.94100	0.94600	0.94478
						WH[g] (30)	0.94100	0.94341	0.94300
Obscuration	205	0.56	0.702	0.20	0.10	PB[e] (29)	0.17	2.37	1.93
						BB[f] (176)	0.72	2.80	1.49
Colour (gold units)	171	5.14	0.336	2.00	0.00	PB[e] (23)	3.00	15.00	7.35
						BB[f] (121)	3.00	23.00	14.32
						WH[g] (27)	9.50	13.00	11.76

[a]N – number of samples; [b]SD – standard deviation; [c]$SEL_{Distell}$ - Distell inter-laboratory Standard Error of Laboratory; [d]$SEL_{own\ data}$ – Standard Error of Laboratory calculated in Chapter 3; [e]PB – potstill brandy; [f]BB – blended brandy; [g]WH – whisky; [h]VO – vodka; [i]GI - gin

Skewness, which is an indication of the lack of normality of statistical distribution of the data poses a challenge for the calibration sample selection. Large skewness (usually bigger than the absolute value of 1) also has implications for calculation of the RPD value, since large skewness affects the value of SD, which, in turn is used for calculation of RPD (Bellon-Maurel *et al*., 2010). The selection of the calibration and validation sets were therefore based on the PCA score plots, taking the variation in all directions on 2-dimensional score plots into consideration. The $SEL_{Distell}$ and $SEL_{own\ data}$ (as calculated in Chapter 3) are reported.

4.3.2 Evaluation of the spectra

Figure 4.1 shows the line plot of all the samples with all the wavenumbers 929.78 to 4996.11 cm^{-1}. The two regions with the distinctive water absorbing spectra, 1454.47 to 1778.54 cm^{-1} and 2839.49 to 3460.63 cm^{-1}, are indicated on the spectra (Nieuwoudt *et al,* 2004; Lachenmeier, 2007). The region between 979.93 and 1450.61 cm^{-1} is known as the fingerprint region, and is known to contain a significant amount of variation related to the absorbance by molecular chemical groups present in alcoholic beverages (Nieuwoudt *et al,* 2004). The regions

below 979.93 cm^{-1} and above 3460.63 cm^{-1} were found to contain little useful information and were trimmed off in further data analyses. No spectra were noted with obvious visual atypicalities in the line plots.

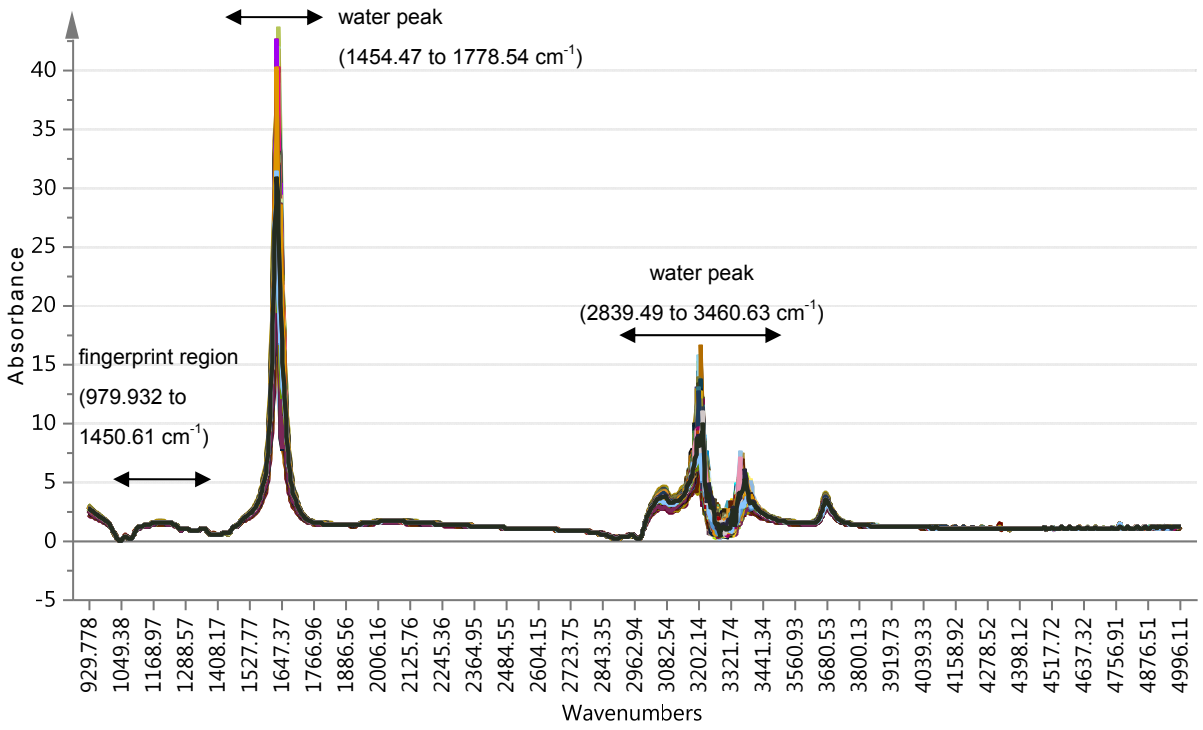


Figure 4.1: Line plot of all the spectra with all wavenumbers 929.778 to 4996.11 cm^{-1}. Water absorbing regions and fingerprint region are indicated.

4.3.3 Explorative data analysis

Although it is common practice to deselect the water absorbing areas in wine spectra (Nieuwoudt *et al*., 2004), it was of interest to investigate the influence of these regions in PCA analysis of high alcohol products, such as spirit products. Figure 4.2 shows the PCA score plot (a) performed on all the samples with all the wavenumbers and corresponding loading plot (b) of PC1 explained 66.8% of the variance in the spectral data. PCA did not show separation between the various product types, except for the potstill brandies that located to the left of PC1 (Figure 4.2 a). The wavenumbers that contributed most to the distribution seen in the scores plot, were those associated with water absorbance, as seen in the loading plot, Figure 4.2 (b). The areas 1454.47 to 1778.54 cm^{-1} and 2839.49 to 3460.63 cm^{-1}, associated with water absorbance were deselected in subsequent analysis. The uninformative wavenumbers 929.78 to 979.93 cm^{-1} and 3460.63 to 5012 cm^{-1}, respectively the two extreme ends of the raw spectra, were also removed from further analyses.

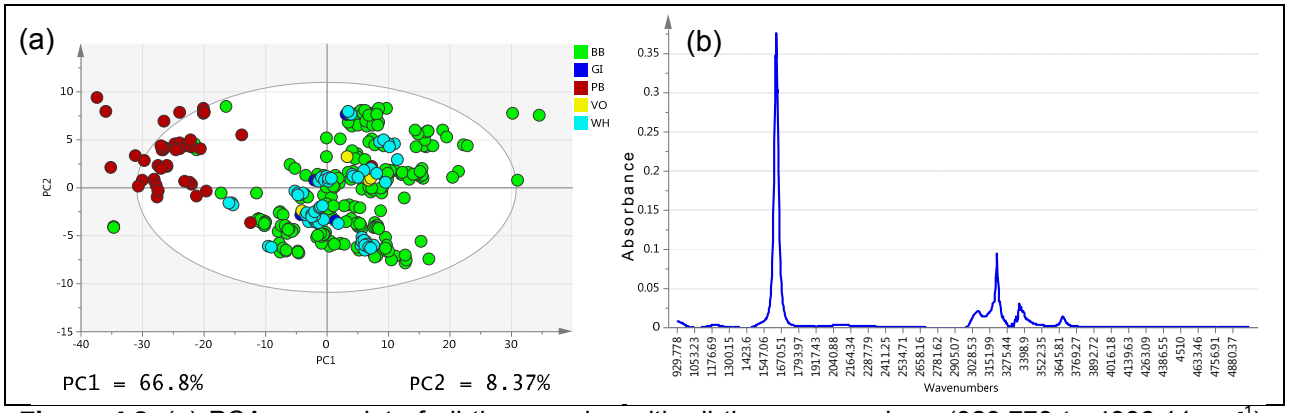


Figure 4.2: (a) PCA score plot of all the samples with all the wavenumbers (929.778 to 4996.11 cm^{-1}). Samples were coloured according the product type, namely potstill brandies (PB) (red markers), blended brandies (BB) (green markers), whiskies (WH) (light blue markers), gins (GI) (dark blue markers) and vodkas (VO) (yellow markers). Ellipse on the score plot indicates Hotelling T^2 range at 95% confidence; (b) Loading plot with PC1 (explained 66.8% of data variance).

Considering the effect of the water regions on the PCA, subsequent PCA was performed with the wavenumbers that contained the most useful information, 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}. The results are shown in Figure 4.3.

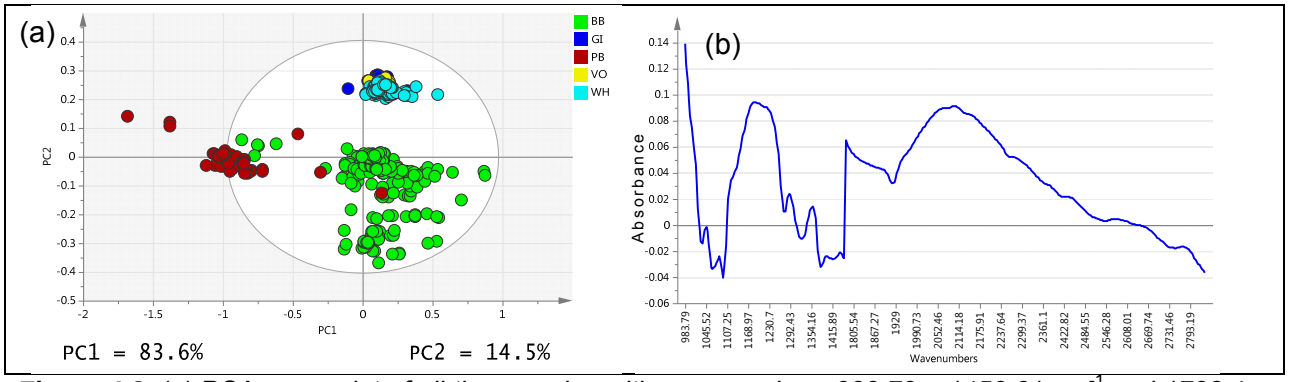


Figure 4.3: (a) PCA score plot of all the samples with wavenumbers 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}. Samples were coloured according the product type, namely potstill brandies (PB) (red markers), blended brandies (BB) (green markers), whiskies (WH) (light blue markers), gins (GI) (dark blue markers) and vodkas (VO) (yellow markers). Ellipse on the score plot indicates Hotelling T^2 range at 95% confidence; (b) Loadings plot of PC1 (83.6% variance explained).

The PCA score plot in Figure 4.3 (a) showed clear groupings in the dataset. Clearly defined groups of samples may benefit with a separate calibration model (Moreira and Santos, 2004). The results also indicated the total explained variance of PC1 and PC2 increased from 75.17% to 98.1%, which is frequently found with spectroscopy data (Esbensen, 2002). High loadings were seen in the fingerprint region of 979.93 to 1450.61 cm^{-1} (Nieuwoudt *et al.,* 2004) after the removal of the water regions, confirming the influence of the chemistry of the samples in the patterns observed in the PCA analysis and also the necessity for the removal of the water absorbing wavenumbers.

As mentioned before, the potstill brandies showed a clear separation from the rest of the samples on PC1 in Figure 4.3 (a). Potstill brandies have an average ethanol concentration of 38 %v/v compared to the average ethanol concentration of 43 %v/v for the rest of the spirit products in this data set (see Table 4.2). Separation on ethanol concentration is confirmed by the interpretation of the loadings plot in Figure 4.3 (b). The high loading at 1080 cm^{-1} is characteristic of the absorption by OH alcoholic groups and have been identified in various studies on various alcoholic beverages from beer to spirits (Gallignani *et al.*, 2005), on

Australian wines (Cozzolino *et al.*, 2009), red and white wines (Bevin *et al.,* 2008). Separation along ethanol concentration of spirit products was also found in a previous study done by Da Costa *et al.* (2004).

The potstill brandy samples were deselected and PCA (Figure 4.4) was done on the remaining data to explore the separation amongst the remaining product types. Two groupings were revealed separating on PC1, one consisting of the blended brandies and the other consisting of the whiskies, vodkas and gins grouped together. The blended brandy grouping on the negative of PC1 showed groupings within the product type, which can be attributed to the differences between different brands of blended brandies used in this study. PC1 and PC2 explain 53.9% and 41.1% of the data variance, respectively. The peaks observed in the loadings plot of PC1 in Figure 4.4 (b) between 999.222 and 1138.11 cm^{-1} indicate the significance of ethanol and sugars on the separation on PC1 between these groups. The importance of these regions has been found in other studies on alcoholic beverages (Cozzolino *et al.,* 2009; Bevin *et al.,* 2008). Similar results were reported by Da Costa *et al*. (2004) who showed separation of samples in PCA analysis, based on their sugar level, where the relatively higher sugar spirits (rum and brandy) were separated from the rest (vodka and whiskey).

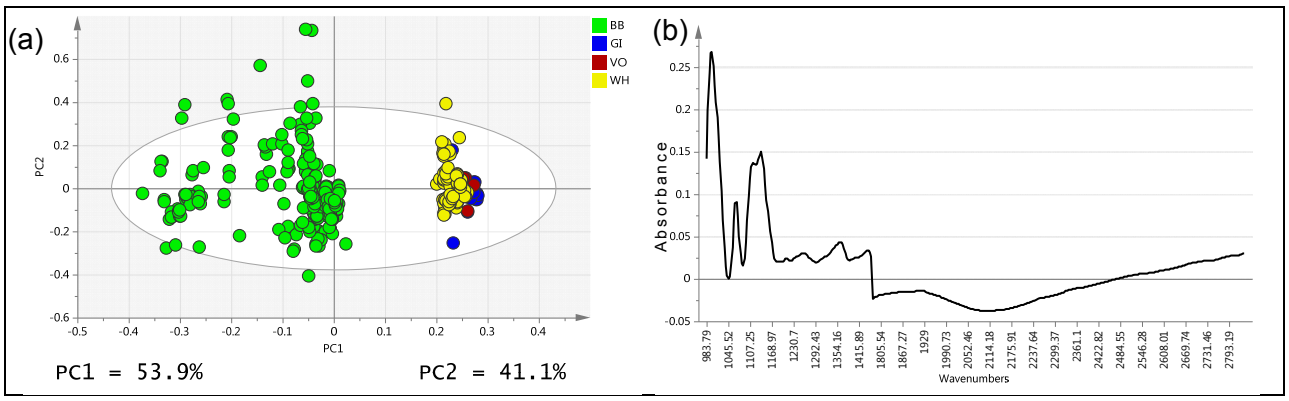


Figure 4.4: (a) PCA score plot of the samples excluding the potstill brandies. Wavenumbers excluding the water absorbing regions (983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}) were used. PC1 (53.9% variance explained) versus PC2 (41.1% variance explained). Samples coloured according to product type, namely blended brandies (BB) (green markers), whiskies (WH) (yellow markers), gins (GI) (dark blue markers) and vodkas (VO) (red markers) are shown on the plot. Ellipse on the score plot indicates Hotelling T^2 range at 95% confidence; (b) Loadings plot of PC1.

Brandies were the only products containing sugar in this data set. The major compounds in alcoholic beverages, such as ethanol and sugar, have shown to affect the quantification of parameters present in lower concentrations and those compounds that have absorption frequencies close to it (Moreira and Santos, 2004). Therefore the effect of other differentiating compounds is dwarfed by the strong presence of sugars in the brandy. This could explain why the whiskies, although vastly different organoleptically and chemically from the clear spirits, did not separate in the PCA score plot in Figure 4.4 (a).

A clear clustering between whiskies and clear spirits (vodkas and gins) was seen when the brandy samples were removed from the PCA (Figure 4.5). Ethanol is the major volatile component in the wine matrix (Cozzolino *et al.,* 2005), concealing effects from minor volatile components. As seen in Figures 4.3 and 4.4, ethanol has the highest influence on PC1. Looking at PCs higher than 1 reveals contribution of minor components in the distilled beverage matrix.

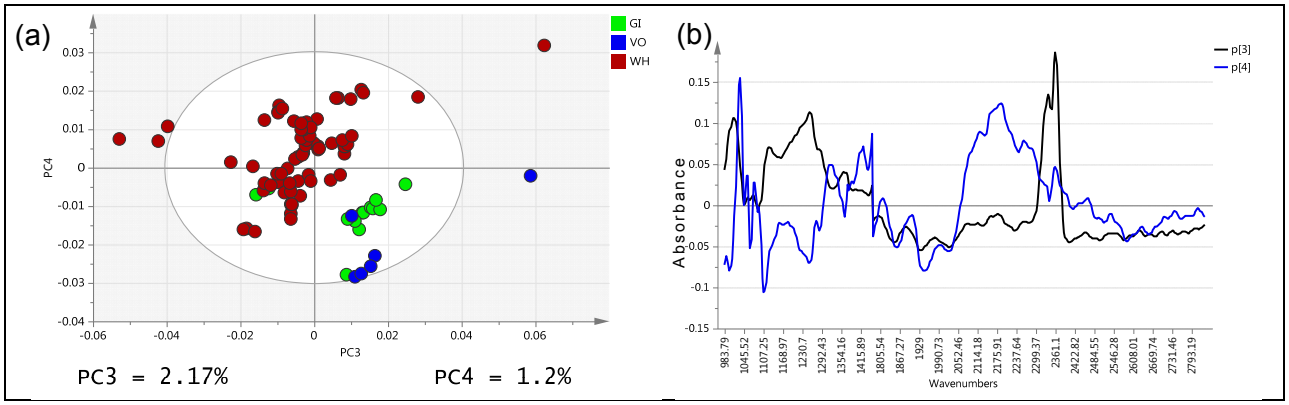


Figure 4.5: a) PCA score plot of the gin, vodka and whisky samples. Wavenumbers excluding the water absorbing regions (983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}) were used. PC3 (2.17% variance explained) versus PC4 (1.2% variance explained). Samples were coloured according the product type, namely whiskies (WH) (red markers), gins (GI) (green markers) and vodkas (VO) (blue markers). Ellipse on the score plot indicates Hotelling T^2 range at 95% confidence; (b) Loadings plot of PC3 and PC4.

PC3 and PC4, respectively, explained 2.17% and 1.2% of the variance in the X-data. . Wavenumbers with high loadings were found around 1454 cm^{-1} and relate to CO=O and C=C, C-H_2, C-H_3 vibrations associated with organic acids and aldehydes (Cozzolino *et al.*, 2009). High loadings were also observed between 1288 and 1458 cm^{-1} and contain information related to C ═ O stretching for aldehydes, carboxylic acids and esters (Cozzolino *et al.*, 2009). From the interpretation of the scores and loading plots, it can be concluded that differences between whiskies and clear products can be partly interpreted on the basis of higher alcohols, aldehydes and esters.

Organoleptically, whisky, vodka and gin products differ greatly. Whiskies are rich in flavourful compounds originating from the raw material (mash or grain) (SA Liquor Products Act 60 of 1989 and aging in barrels, while gin is infused with juniper berries and vodka is flavourless. These clear groupings amongst product types imply that the approach of investigating the development of separate calibration models for each product is justified.

PCA of the samples based on the **y**-variables showed the relationship between the **y**-variables and the positioning of the samples on the score plot. The score plot results in Figure 4.6 (a) were coloured according to product type. The loadings plot in Figure 4.6 (b) indicated ethanol to have a high positive loading on PC1. Figure 4.6 (a) indicated the ethanol concentration increasing from left to right on PC1, in accordance with the loading plot and results of Figure 4.3 (a). Density had a high negative loading on PC1, confirming the high density values in potstill brandies (Table 4.2). Obscuration with the high positive loading on PC2 indicate the samples on the bottom half of the score plot in Figure 4.6 (a) to have the least sugar in them, supporting the positioning of the whiskies, gins and vodkas on the score plot at the bottom half of PC2. These products contain the least amounts of sugar in them. Colour has a high positive loading on PC2 supporting the positioning of the clear spirits at the bottom half of PC2 in Figure 4.6 (a).

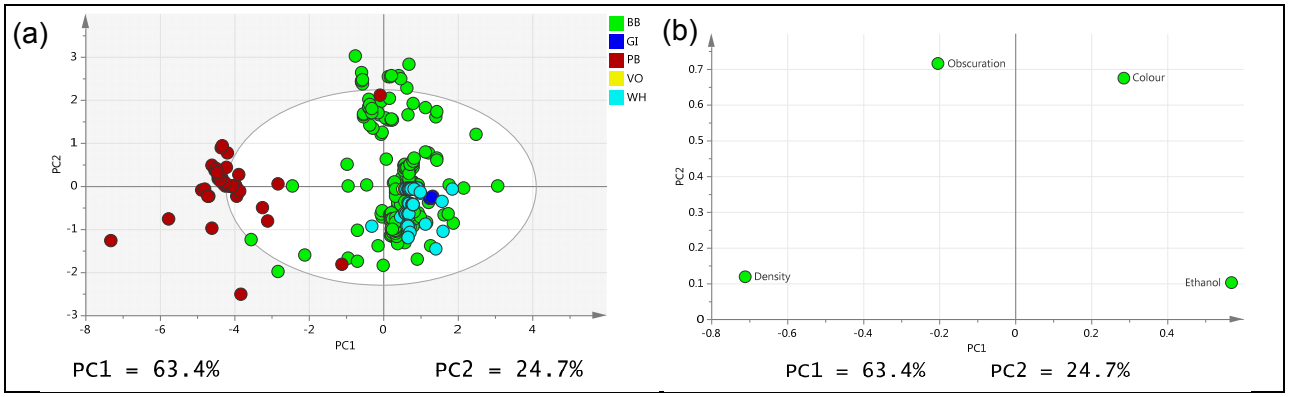


Figure 4.6: PCA performed on all the samples and **x**- and **y**-variables. The wavenumbers 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1} were used. All four parameters ethanol, density, obscuration and colour were included. (a) Score plot was coloured according to product type. Potstill brandies (PB) (red markers), blended brandies (BB) (green markers), whiskies (WH) (light blue markers), gins (GI) (dark blue markers) and vodkas (VO) (yellow markers) are indicated in the plot. Ellipse on the score plot indicates Hotelling T^2 range at 95%confidence.. (b) Loading plot of the four **y**-variables.

4.3.4 PLS regression modelling

For each of the four parameters (ethanol, density, obscuration and gold), PLS calibration models were developed using the whole dataset and separated on product types based on the PCA results described in section 4.3.3. Three strategies were used for the development of PLS models, as described in section 4.2.4.2.

The number of components per calibration was selected based on the highest explained variance with the lowest number of components to combat overfitting whilst sufficiently modelling the variance in the data set.

For the wavenumber selection procedure in *Strategy 3,* the spectra of the samples with the highest and lowest values for each parameter were weighed up against each other to identify the wavenumbers of significance for the specific parameter. The plot for ethanol wavenumber selection is shown in Figure 4.7. The peaks at which the two spectra diverged were chosen. The goal of wavenumber selection is to identify a subset of spectral frequencies that produce the smallest possible errors when making quantitative determinations or discriminating between dissimilar samples (Balabin and Smirnov, 2011). The selection removes uninformative variables, resulting in an uncluttered model with better prediction results. The wavenumber selection was only tested on the second derivative S.G. spectra.

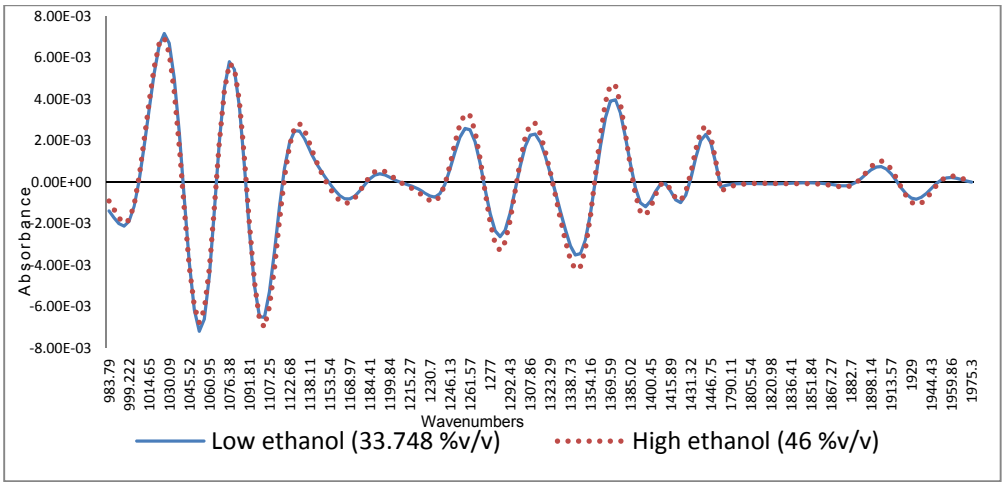


Figure 4.7: Second derivative S.G. FT-IR spectra (983.79 – 1450.61 cm^{-1} and 1782.4 – 1975.3 cm^{-1}) of a sample with a low ethanol value (33 %v/v) and a sample of a high ethanol value (46 %v/v).

These peaks indicated in the high-low plots, the wavenumbers indicated as significant in the PCAs score plot separation of the products and the wavenumbers with the highest loading weights were initially chosen during the wavenumber selection process per calibration for the third strategy to variables in the calibration models. The RMSEP is monitored throughout the selection process. Wavenumbers selected for *Strategy 3* for each parameter is given in Table 4.3.

Table 4.3: Wavenumbers selected for *Strategy 3* for each of the four parameters.

Ethanol	Density	Obscuration	Colour
		983.79	983.79
		987.65	987.65
		991.51	991.51
		995.36	995.36
		999.22	999.22
	1018.51		
		1026.23	1026.23
1045.52			
1053.23		1053.23	1053.23
1080.24		1076.38	1076.38
1103.39	1103.39		1103.39
1130.39		1130.39	1130.39
1165.12		1165.12	1165.12
1192.12		1234.56	1234.56
1234.56	1242.28		
1257.71			
1284.71			
1311.72	1304.00		
1342.58			
1358.02			
1369.59			
1396.60			
1442.89		1450.61	1450.61
	1901.99		
1905.85			
1932.86	1932.86		
	2067.89		
	2145.05		
	2391.96		
	2430.54		
	2723.75		
	2758.47		
	2800.91		

The regression statistics for the calibration model with the lowest RMSEP and highest R^2 was reported. Comparative results of the different calibration models are shown in Table 4.4. The calibration sets with too little samples were not reported.

Table 4.4: Comparison of regression statistics for the various calibration models.

Parameters and SEL	Sample sets	*Strategy 1*					*Strategy 2*					*Strategy 3*				
		No of samples (calibration set : validation set)	RMSECV [c]	R^2 [d]	RMSEP [e]	RPD [f]	No of samples (calibration set : validation set)	RMSECV [c]	R^2 [d]	RMSEP [e]	RPD [f]	No of samples (calibration set : validation set)	RMSECV [c]	R^2 [d]	RMSEP [e]	RPD [f]
ETHANOL $SEL_{Distell}$ [a] = 0.200 %v/v $SEL_{own\ data}$ [b] = 0.097 %v/v	All products	238 : 117	0.425 %v/v	0.99	0.314 %v/v	5.66	231:114	0.275 %v/v	0.99	0.291 %v/v	5.87	229 :95	0.297 %v/v	0.99	0.246 %v/v	6.97
	All brandy	176 : 91	0.463 %v/v	0.98	0.351 %v/v	5.68	179 : 78	0.354 %v/v	0.99	0.270 %v/v	7.63	140 : 66	0.254 %v/v	0.99	0.254 %v/v	8.21
	Blended brandy	130 : 67	0.145 %v/v	0.76	0.172 %v/v	1.36	135 : 68	0.149 %v/v	0.80	0.153 %v/v	1.66	125 : 56	0.158 %v/v	0.94	0.106 %v/v	2.87
	Potstill brandy	29 : 15	0.525 %v/v	0.97	0.397 %v/v	3.53	26 : 15	0.055 %v/v	0.96	0.196 %v/v	4.60	19 : 10	0.052 %v/v	0.96	0.067 %v/v	2.67
	Whisky & clear spirits	58 : 26	0.077 %v/v	0.94	0.063 %v/v	2.41	57 : 26	0.082 %v/v	0.86	0.070 %v/v	2.14	50 : 28	0.062 %v/v	0.86	0.078 %v/v	1.99
	Whisky	46 : 18	0.060 %v/v	0.88	0.077 %v/v	2.03	46 : 20	0.078 %v/v	0.89	0.072 %v/v	2.14	46 : 20	0.073 %v/v	0.92	0.089 %v/v	1.73
	Clear spirits	9 : 6	0.057 %v/v	0.92	0.039 %v/v	2.06	10 : 5	0.028 %v/v	0.92	0.035 %v/v	2.22	13 : 4	0.059 %v/v	0.99	0.038 %v/v	2.30
DENSITY $SEL_{Distell}$ [a] = 0.00100 $^{20}_{20}$ $SEL_{own\ data}$ [b] = 0.00360 $^{20}_{20}$ (clear spirits) and 0.00031 $^{20}_{20}$ (brandies)	All products	129 : 65	0.00041 $^{20}_{20}$	0.96	0.00064 $^{20}_{20}$	4.49	129 : 65	0.00053 $^{20}_{20}$	0.98	0.00057 $^{20}_{20}$	5.05	129 : 63	0.00067 $^{20}_{20}$	0.98	0.00053 $^{20}_{20}$	4.98
	All brandy	103 : 50	0.00058 $^{20}_{20}$	0.98	0.00054 $^{20}_{20}$	5.16	105 : 52	0.00048 $^{20}_{20}$	0.98	0.00049 $^{20}_{20}$	5.81	117 : 41	0.00062 $^{20}_{20}$	0.96	0.00065 $^{20}_{20}$	3.41
	Blended brandy	77 : 39	0.00045 $^{20}_{20}$	0.94	0.00031 $^{20}_{20}$	2.78	87 : 40	0.00034 $^{20}_{20}$	0.83	0.00047 $^{20}_{20}$	1.88	92 : 41	0.00042 $^{20}_{20}$	0.86	0.00039 $^{20}_{20}$	1.97
	Whisky & clear spirits	24 : 9	0.00019 $^{20}_{20}$	0.99	0.00010 $^{20}_{20}$	5.86	20 :10	0.00005 $^{20}_{20}$	0.96	0.00011 $^{20}_{20}$	3.45	23 : 8	0.00016 $^{20}_{20}$	0.92	0.00017 $^{20}_{20}$	2.59
	Whisky	20 : 10	0.00014 $^{20}_{20}$	0.96	0.00021 $^{20}_{20}$	2.57	20 : 10	0.00027 $^{20}_{20}$	0.99	0.00015 $^{20}_{20}$	3.54	20 : 9	0.00018 $^{20}_{20}$	0.94	0.00025 $^{20}_{20}$	3.05

Parameters and SEL	Sample sets	*Strategy 1*					*Strategy 2*					*Strategy 3*				
		No of samples (calibration set : validation set)	RMSECV [c]	R^2 [d]	RMSEP [e]	RPD [f]	No of samples (calibration set : validation set)	RMSECV [c]	R^2 [d]	RMSEP [e]	RPD [f]	No of samples (calibration set : validation set)	RMSECV [c]	R^2 [d]	RMSEP [e]	RPD [f]
OBSCURATION $SEL_{Distell}$ [a] = 0.20 $SEL_{own\ data}$ [b] = 0.10	All brandy	136 : 63	0.12	0.97	0.13	4.10	136 : 68	0.18	0.96	0.17	3.33	134 : 67	0.16	0.97	0.13	4.47
	Blended brandy	117 : 56	0.12	0.97	0.12	4.46	114 : 58	0.14	0.97	0.13	4.16	122 : 41	0.12	0.98	0.10	5.26
	Potstill brandy	19 : 9	0.06	0.99	0.09	4.91	16 : 10	0.05	0.97	0.07	5.35	18 : 8	0.09	0.98	0.09	4.57
COLOUR $SEL_{Distell}$ [a] = 2.00 gold units $SEL_{own\ data}$ [b] = 0 gold units	All products	115 : 58	1.31 gold units	0.95	1.52 gold units	3.35	110 : 53	2.11 gold units	0.90	2.10 gold units	2.36	98 : 49	2.35 gold units	0.81	2.29 gold units	1.99
	All brandy	96 : 48	0.89 gold units	0.96	1.62 gold units	3.42	91 : 46	0.96 gold units	0.96	1.63 gold units	3.50	81 : 41	1.63 gold units	0.93	2.25 gold units	2.27
	Blended brandy	80 : 40	0.94 gold units	0.96	1.48 gold units	3.51	81 : 36	1.07 gold units	0.97	1.14 gold units	4.43	75 : 25	1.55 gold units	0.95	1.33 gold units	3.68
	Whisky	17 : 8	0.35 gold units	0.78	0.52 gold units	1.67	17 : 6	0.23 gold units	0.89	0.45 gold units	2.27	19 : 9	0.72 gold units	0.73	0.76 gold units	1.31

[a] $SEL_{Distell}$ – Standard error of laboratory for a Distell laboratory; [b] $SEL_{own\ data}$ – Standard error of laboratory from validation study (see Chapter 3); [c] RMSECV – Root mean squared error of cross validation; [d] R^2 – correlation coefficient; [e] RMSEP – residual mean standard error of prediction; [f] RPD – Residual predictive deviation

4.3.4.1 Ethanol PLS regression models

Strategy 3 proved the best strategy according to the results in Table 4.4. This could be due to the removal of uninformative wavenumbers shielding the effects of the valuable wavenumbers. The predicted versus measured plot shows the spread of the samples and the fit of the linear regression line (Figure 4.8). The gap between 39 and 43 %v/v was expected due to the commercial nature of this study, and is not ideal.

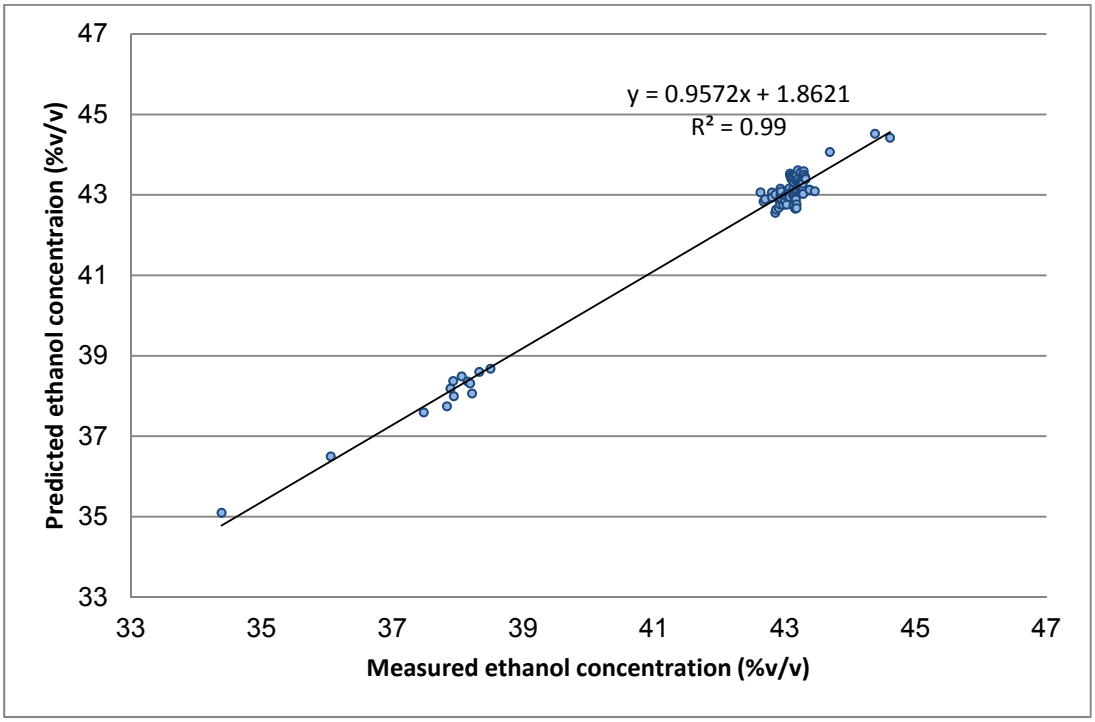


Figure 4.8: Regression plot of predicted versus measured ethanol concentration for the brandy calibration for ethanol concentration (%v/v).

Splitting the dataset into product type, according the PCA results in Figures 4.3 (a) and Figure 4.5 (a), yielded more accurate results when considering the RMSEP values, which were comparable to the $SEL_{Distell}$ of 0.200 %v/v. The calibration models for potstill brandy (RMSEP = 0.067 %v/v), whisky and clear spirits combined (RMSEP = 0.078 %v/v), whisky only (RMSEP = 0.089 %v/v) and clear spirits only (RMSEP = 0.038 %v/v), out-performed the laboratory method ($SEL_{own\ data}$ of 0.097 %v/v) in terms of accuracies achieved. Although this was a feasibility study, in contradiction to the generally held notion, that predicted values will always be less accurate than the reference data, the results obtained in this study, clearly showed the contrary. All the reported RMSEP values were well below labeling tolerances of 0.5 %v/v (SA Liquor Products Act 60 of 1989), which allows a discrepancy of ±0.5 %v/v between the ethanol concentration stated on the label and the actual analyses results. The RMSEP values obtained with the product specific calibration models in this study, showed a significant improvement on reports in the literature where FT-IR spectroscopy was used for prediction in spirit products. Lachenmeier (2007) reported a RMSEP value of 0.21 %v/v and R^2 of 0.94. Gallignani *et al.* (1994) reported a R^2 of 0.99 and average prediction error of 0.20 %v/v, in a strategy that required prior dilution of samples with ethanol concentration higher than 15 %v/v. McIntyre *et al.* (2011) determined ethanol concentrations in the range of 35 – 45 %v/v with an accuracy of 0.8% with PLS regression.

RPD values for the calibration models including respectively all the samples (6.97) and all brandies (8.21) indicated that the models developed were suitable for quantification (Pink *et al*., 1998). RPD values of the other product specific calibrations, ranging between 1.73 and 2.87, indicate these models unsuitable for accurate quantification (Table 4.1) (Pink *et al.* 1998). The low RPD values for the product specific calibrations can be ascribed to the narrow range of reference values and subsequently lower SD. If the error is large compared to the range and hence of SD of the reference data, a small RPD value results, and the calibration model is considered not robust. The higher the value of the RPD, the greater the probability of the model to predict the chemical composition accurately in samples outside the calibration set (Cozzolino *et al*., 2008). Including more samples for all the products will extend the range and produce a model with possibly larger SD, thereby resulting in a higher RPD, which can be rendered suitable for quantification.

R^2 indicated excellent precision R^2 = 0.92 – 0.99, except for the calibration model with the whisky and clear spirits combined, achieving good precision with R^2 = 0.856, according to the interpretation guidelines of Urbano-Cuadrado *et al*. (2005).

4.3.4.2 Density PLS regression models

Strategy 1 and *Strategy 2* yielded the best calibration statistics, as shown in Table 4.4. The predicted versus measured plot of all the samples without pre-processing or wavenumber selection is given in Figure 4.9. The gap between 0.98400 and 0.95400 ($^{20}_{20}$). was expected due to the commercial nature of this study, and is not ideal.

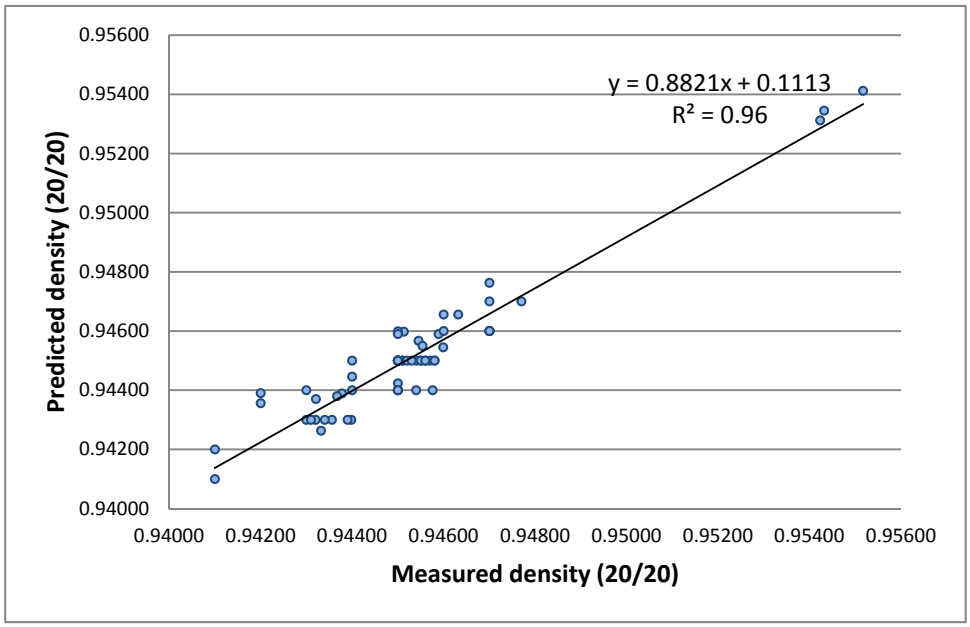


Figure 4.9: Regression plot of predicted versus measured density for all the samples without any pre-processing or wavenumber selection.

All of the calibrations surpassed the goal of the $SEL_{Distell}$ of 0.00100 ($^{20}_{20}$). The RMSEP for the calibrations that included all the samples (0.00064 ($^{20}_{20}$), 0.00057 ($^{20}_{20}$)) compared well with a previous report. Lachenmeier (2007) reported a SECV of 0.00070 ($^{20}_{20}$) for spirit products. All the product specific calibration models gave more precise results than the calibration model based on all product types according to the RMSEP values presented here for models that used *Strategy 1* and *Strategy 2*, respectively; blended brandies = 0.00031 ($^{20}_{20}$); 0.00047 ($^{20}_{20}$), whisky

and clear spirits combined = 0.00010 ($^{20}_{20}$); 0.00011 ($^{20}_{20}$) and whisky = 0.00021 ($^{20}_{20}$); 0.00015 ($^{20}_{20}$). The advantage of developing separate models for the different products to increase the precision was again shown in the results obtained with the density calibrations. Calibration models for whisky only and whisky and clear spirits combined, out-performed the laboratory error ($SEL_{own\ data}$ of 0.00031 ($^{20}_{20}$)).

When evaluating the RPD values, *Strategy 1* calibration models for brandies (5.16) and whisky and clear spirits combined (5.86) proved suitable for screening purposes (Table 4.1). Calibration models based on all products (4.49), blended brandies (2.78) and whisky (2.57), were suitable for screening (Pink *et al.,* 1998). In the case of *Strategy 2* the calibration models with all the products (5.05) and brandies (5.81) proved suitable for quantification. The whisky only model (3.54) and whisky and clear spirits combined calibration model (3.45) proved suitable for screening, while the blended brandy (1.88) model was unsuited for quantification.

R^2 values ranged from 0.94 to 0.99 for the *Strategy 1* calibration models, indicating excellent precision. *Strategy 2* calibrations also achieved excellent accuracy with R^2 ranging from 0.96 to 0.99, except for the blended brandy (0.83), yielding good precision, rendering this model unsuited for quantification, based on the interpretation guidelines of Urbano-Cuadrado *et al*. (2005).

4.3.4.3 Obscuration PLS regression models

As previously noted, only brandy has an obscuration value, thus only variations of calibrations with the brandy samples were done. *Strategy 3* yielded the best performing calibrations, as indicated by the. RMSEP value of 0.13 (Table 4.4.) for the calibration with both types of brandies combined, compared well with the $SEL_{Distell}$ of 0.20. The predicted versus measured plot is shown in Figure 4.10. The regression plot shows a lack of samples between 1.5 and 1.9, this can be accounted due to the commercial nature of this study, only using the samples sent in for routine analyses.

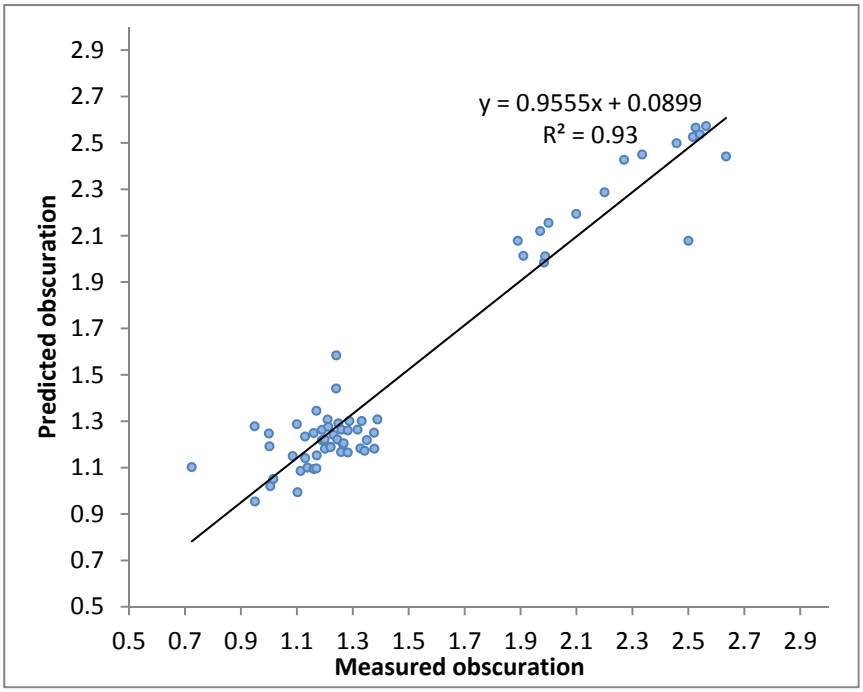


Figure 4.10: Regression plot of predicted versus measured obscuration for all the brandy samples with second derivative S.G. pre-processing with wavenumber selection.

Calibration models for respectively blended brandies (RMSEP = 0.10) and potstill brandies (RMSEP = 0.09), out-performed the $SEL_{own\ data}$ of 0.10. This verified the findings in the ethanol and density calibration results that minimising the variation between the samples lowered the error of the respective calibration models.

The RPD values for the calibration combining the two types of brandies (4.47) and the calibration model for the potstill brandies (4.57) just fell short of the quantification suitability value of RPD of 5 (Table 4.1) (Pink *et al.,* 1998). The calibration model for the blended brandies achieved a RPD of 5.26 indicated the model is suitable for quantification.

The R^2 of all three calibration models were higher than 0.90 and indicated excellent precision (Urbano-Cuadrado *et al*., 2005).

To the best knowledge of the authors, no other studies have reported on the obscuration quantification of spirit products. Since the obscuration is correlated to degree of sweetness, it was compared to studies that reported on sugar quantification in alcoholic beverages. Urbano-Cuadrado *et al.* (2004) achieved a SECV of 0.27 g/L and R^2 of 0.71 in the determination of residual sugar in wines, using Near-infrared spectroscopy. The calibration models in this study showed a vast improvement on the results of the reported study.

4.3.4.4 Colour PLS regression models

In the calibration models for colour, *Strategy 1* and *Strategy 2* proved better performing (see results in Table 4.4). The plot for predicted versus measured for all the samples without pre-processing or wavenumbers selected is shown in Figure 4.11.

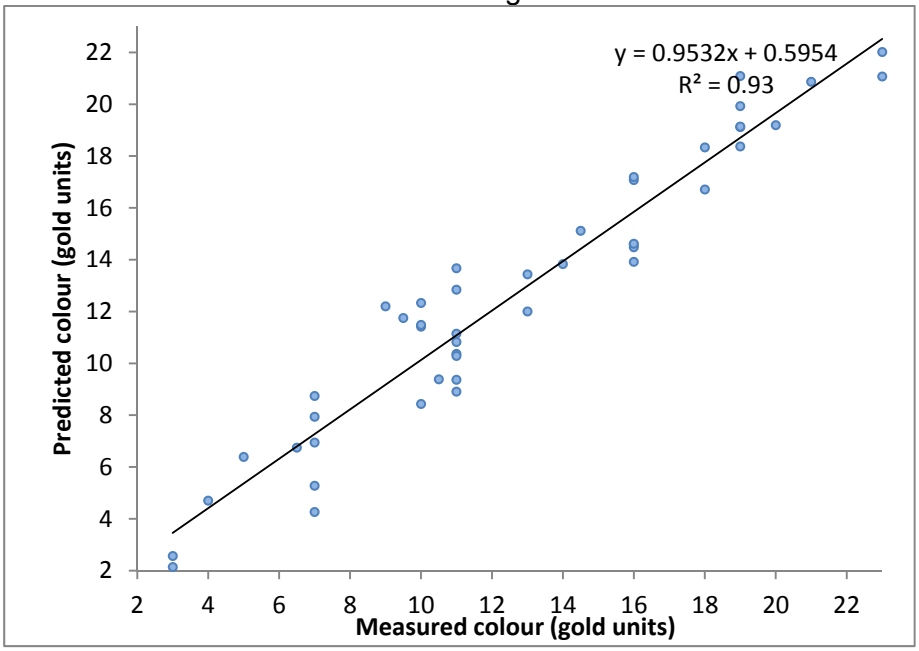


Figure 4.11: Regression plot of predicted versus measured colour for all the samples without any pre-processing or wavenumber selection.

All of the calibrations were in-line with or better performing than the $SEL_{Distell}$ of 2.00 gold units, with RMSEP values ranging from 0.45 to 2.1 gold units. The $SEL_{own\ data}$ of 0 gold units was out of reach. This could be due to the variations included in the calibration data set, not present in the validation study. Variations such as different analysts and different laboratories, were included when the calibration data was accumulated; taking into account the subjectivity of the reference method.

The RPD for these models indicated screening suitability (3<RPD<5), except for the *Strategy 2* model for all the products (2.36) and *Strategy 1* and *Strategy 2* whisky models (1.67 and 2.27) which indicated these models unsuitable for quantification (Pink *et al*., 1998).

The R^2 > 0.9 indicates excellent precision for the above mentioned models that were suitable for screening, while the models not suitable for quantification showed good precision (Urbano-Cuadrado *et al*., 2005).

Correlations between sensory qualities and objective instrumental measurements have been reported in the wine industry, with the focus on flavour compounds (Cozzolino *et al*., 2005, 2006). Wine colour has been successfully analysed with infrared spectroscopy by Urbano-Cuadrado *et al*. (2004), presented acceptable results for colour in red wines SEP = 1.83. McIntyre *et al.* (2011) analysed dried residue of whisky and caramel and using PCA, distinguished between different caramel colourants and different whisky samples, but no quantification of colourants were done.

4.4 CONCLUSION

The aim of this study was to evaluate FT-IR spectroscopy as a rapid means of analysis for spirit products for ethanol, density, obscuration and colour. The performance criteria used to evaluate the calibration models were RMSEP, R^2 and RPD. Test set validation indicated that the models were accurate and robust. Overall, the results of this study suggest that this technique can therefore be used for quantification and screening purposes of spirit products.

In this feasibility study in some cases the RMSEP proved better results than the RMSECV, in contradiction to widely accepted statistics. In further testing this will be verified. It should be highlighted that the sample distribution was not continuous nor optimal, but a true reflection of the commercial nature of this study, evident from the regression plots.

Although not all the calibration models were suitable for quantification, some models were suitable for screening and will save a lot of time and money. The samples can be screened and the borderline or questionable samples can be analysed by the reference method. Future efforts will aim to increase the samples per product type, so to extend the SD within a product type data set to yield an acceptable RPD of more than five to prove the suitability to quantification. The colour data set had the biggest range which resulted in a high SD, but only two product types, brandies and whiskies, were analysed in this model, since the other samples were clear.

Most of the models proved excellent precision with R^2 higher than 0.9. These regression statistics indicated that minimising the variation in a sample set, increases the accuracy of the calibration model.

The technique proposed can greatly increase the efficiency of laboratory operations and throughput of spirit product analyses. These results confirm the versatility and accuracy of infrared instruments in the analysis of these four parameters for spirit products.

This methodology can also be used as a method to classify type of spirit product. This study is a first attempt at classification of South African spirit products. Previous studies have included South African spirits, but for classification of country of origin (Palma and Barroso, 2002). The interpretation of the loadings for the PCAs indicated the spectral differences observed in the products might be caused by differences in organic components such as ethanol content, sugars as well as compounds containing aromatic rings and organic acids, esters and aldehydes. Possible future research strategies could include the analysis of guaranteed authenticate reference samples.

4.5 REFERENCES

1. Association of Official Analytical Chemists (AOAC). (2005). Official methods of analysis of the Association of Official Analytical Chemists, *Association of Official Analytical Chemists*, Arlington, VA.
2. Balabin, R. M., & Smirnov, S.V. (2011). Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel data. *Analytica Chimica Acta.* **692**, 63-72.
3. Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.-M. & McBratney, A. (2010). Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *Trends in Analytical Chemistry* **29,** 1073 – 1081.
4. Bevin, C.J., Dambergs, R.G., Fergusson, A.J. & Cozzolino, D. (2008). Varietal discrimination of Australian wines by means of mid-infrared spectroscopy and multivariate analysis. *Analytica Chimica Acta* **621,** 19-23.
5. Cen, H. & He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science & Technology.* **18**, 72-83.
6. Cozzolino, D., Holdstock, M., Dambergs, R.G., Cynkar, W.U. & Smith, P.A. (2009). Mid infrared spectroscopy and multivariate analysis: A tool to discriminate between organic and non-organic wines grown in Australia. *Food Chemistry* **116,** 761-765.
7. Cozzolino, D., Kwiatkowski, M.J., Dambergs, R.G., Cynkar, W.U., Janik, L.J, Skouroumounis, G. & Gishen, M. (2008). Analysis of elements in wine using near infrared spectroscopy and partial least squares regression. *Talanta* **74**, 711–716.
8. Cozzolino, D., Kwiatkowski, M.J., Parker, M., Cynkar, W.U., Dambergs, R.G., Gishen, M. & Herderich, M.J. (2004). Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Analytica Chimica Acta,* **513** (1) 73-80.
9. Cozzolino, D., Liu, L., Cynkar, W.U., Dambergs, R.G., Janik, L., Colby, C.B. & Gishen, M. (2007). Effect of temperature variation on the visible and near infrared spectra of wine and the consequences on the partial least square calibrations developed to measure chemical composition. *Analytica Chimica Acta* **588,** 224-230.
10. Cozzolino, D., Smyth, H.E., Lattey, K.A., Cynkar, W., Janik, L., Dambergs, R.G., Francis, I.L. & Gishen, M. (2005). Relationship between sensory analysis and near infrared spectroscopy in Australian Riesling and Chardonnay wines. *Analytica Chimica Acta* **539,** 341-348.
11. Cozzolino, D., Smyth, H.E., Lattey, K.A., Cynkar, W., Janik, L., Dambergs, R.G., Francis, I.L. & Gishen, M. (2006). Combining mass spectrometry based electronic nose, visible-near infrared spectroscopy and chemometrics to assess the sensory properties of Australian Riesling wines. *Analytica Chimica Acta* **563,** 319-324.
12. Cozzolino, D. (2012). Recent Trends on the Use of Infrared Spectroscopy to Trace and Authenticate Natural and Agricultural Food Products. *Applied Spectroscopy Reviews,* **47**, 518-530.
13. Craig, A., Cloarec, O., Holmes, E., Nicholson, J.K. & Lindon, J.C. (2006). Scaling and Normalization Effects in NMR Spectroscopic Metabonomic Data Sets. *Analytical Chemistry* **78**, 2262 – 2267.
14. Da Costa R.S., Santos S.R.B., Almeida L.F., Nascimento E.C.L., Pontes M.J.C., Lima R.A.C., Simões S.S. & Araújo M.C.U. (2004). A novel strategy to verification of adulteration in alcoholic beverages based on Schlieren effect measurements and chemometric techniques. *Microchemical Journal,* **78,** 27– 33.
15. Di Egidio, V., Oliveri, P., Woodcock, T. & Downey, G. (2011). Confirmation of brand identity in foods by near infrared transflectance spectroscopy using classification and class-modelling chemometric techniques — The example of a Belgian beer. *Food Research International* **44,** 544-549.

16. Esbensen, K.H., (2002). Multivariate Data Analysis – In Practice. An introduction to multivariate data analysis and experimental design, *5th ed*. Camo ASA, Oslo.
17. Gallignani, M., Ayala, C., Del Rosario Brunetto, M., Burguera, J.L. & Burguera, M. (2005). A simple strategy for determining ethanol in all types of alcoholic beverages based on its on-line liquid-liquid extraction with chloroform, using a flow injection system and Fourier transform infrared spectrometric detection in the mid-IR. *Talanta* **68,** 470-479.
18. Gallignani, M., Garrigues, S. & De la Guardia, M. (1994). Derivative Fourier transform infrared spectrometric determination of ethanol in alcoholic beverages. *Analytica Chimica Acta*, **287,** 275-283.
19. Lachenmeier, D.W. (2007). Rapid quality control of spirit drinks and beer using multivariate data analysis of Fourier transform infrared spectra. *Food Chemistry,* **101**, 825-832.
20. McIntyre, A.C., Bilyk, M.L., Nordon, A., Golquhoun, G. & Littlejohn, D. (2011). Detection of counterfit Scothc whisky samples using mid-infrared spectrometry with an attenuated total reflectance probe incorporating polycrystalline silver halide fibres. *Analytica Chimica Acta* **690,** 228-233.
21. Moreira, J.L. & Santos, L. (2004). Spectroscopic interferences in Fourier transform infrared wine analysis. *Analytica Chimica Acta* **513,** 263-268.
22. Næs, T., Isaksson, T., Fearn, T. & Davies, T. (2004). A User-Friendly Guide to Multivariate Calibration and Classification. NIR Publications, Chichester, United Kingdom.
23. Nieuwoudt, H.H., Pretorius, I.S., Bauer, F.F., Nel, D.G. & Prior, B.A. (2006). Rapid screening of the fermentation profiles of wine yeasts by Fourier transform infrared spectroscopy. *Journal of Microbiological Methods* **67,** 248-256.
24. Nieuwoudt, H.H., Prior, B.A., Pretorius, I.S., Manley, M. & Bauer, F.F. (2004). Principal Component Analysis Applied to Fourier Transform Infrared Spectroscopy for the Design of Calibration Sets for Glycerol Prediction Models in Wine and for the Detection and Classification of Outlier Samples. *Journal of Agricultural and Food Chemistry* **52**, 3726 - 3735
25. Palma, M. & Barroso, C.G. (2002). Application of FT-IR spectroscopy to the characterisation and classification of wines, brandies and other distilled drinks. *Talanta* **58,** 265-271.
26. Patz, C.-D., Blieke, A., Ristow, R. & Dietrich, H. (2004). Application of FT-MIR spectrometry in wine analysis. *Analytica Chimica Acta* **513**, 81-89.
27. Pink, J., Naczk, M. & Pink, D. (1998). Evaluation of the Quality of Frozen Minced Red Hake: Use of Fourier Transform Infrared Spectroscopy. *Journal of Agriculture and Food Chemistry*, **46,** 3667-3672.
28. Pontes, M.J.C, Santos, S.R.B., Araújo, M.C.U., Almeida, L.F., Lima, R.A.C., Gaião, E.N. & Souto, U.T.C.P. (2006). Classification of distilled alcoholic beverages and verification of adulteration by near infrared spectrometry. *Food Research International* **39,** 182-189.
29. Romera-Fernández, M., Berrueta, L.A., Garmón-Lobato, S., Gallo, B., Vicente, F. & Moreda, J.M. (2012). Feasibility study of FT-MIR spectroscopy and PLS-R for the fast determination of anthocyanins in wine. *Talanta* **88***,* 303-310.
30. Skoog, D.A., Holler, F.J. & Nieman, T.A. (1998). *Principles of Instrumental Analysis*. 5th Edition, Harcourt Brace Company, Florida, USA.
31. Soriano, A., Pèrez-Juan, P.M., Vicario, A., González, J.M. & Pèrez-Coello, M.S. (2007). Determination of anthocyanins in red wine using newly developed method based on Fourier transform infrared spectroscopy. *Food Chemistry* **104,** 1295-1303.
32. South African Liquor Products Act No. 60 of 1989.
33. Urbano-Cuadrado, M., Luque de Castro, M.D., Pérez-Juan, P.M. & Gómez-Nieto, M.A. (2005). Comparison and joint use of near infrared spectroscopy and Fourier transform mid infrared spectroscopy for the determination of wine parameters. *Talanta,* **66**, 218-224.

34. Urbano-Cuadrado, M., Luque de Castro, M.D., Pérez-Juan, P.M., García-Olmo, J. & Gómez-Nieto, M.A. (2004). Near infrared reflectance spectroscopy and multivariate analysis in enology. Determination or screening of fifteen parameters in different types of wines. *Analytica Chimica Acta* **527,** 81-88.

35. Silverstein, R.M., Bassler, G.C. & Morrill, T.C. (1991). Spectrometric Identification of Organic Compounds, 5th edition, Wiley, New York.

Chapter 5

RESEARCH RESULTS

Feasibility study of classification of brandy products using Fourier transform infrared (FT-IR) spectroscopy and multivariate data analysis

CHAPTER 5 – Feasibility study of classification of brandy products using Fourier transform infrared (FT-IR) spectroscopy and multivariate data analysis

ABSTRACT

Product monitoring aimed at protection of the integrity and maintenance of the quality of a brand, as well as assuring authenticity of a brand, has become a priority of producers, with the increase in consumer demands for exposure of fraudulent activity. Classification, and ultimately authentication of brandy products is of extreme importance for the South African (SA) beverage industry, since it is the biggest exported spirit, which competes on an international market. Currently, techniques involved in product authentication, involve lengthy multi-component analyses and comparison of data with previously recorded analyses of authentic samples. This task has not been addressed in full by the SA industry, and here we report a first step towards this end, by establishment of a rapid screening test for classification of SA brandy products. For this purpose, Fourier transform infrared (FT-IR) spectra of potstill and blended brandies were analysed by multivariate data techniques, to evaluate how well the different products could be distinguished from one another, using FT-IR spectra only. The sample set used, consisted of 259 different SA brandy samples of nine different brands and spectra were collected in the wavenumber range 926 to 5012 to cm^{-1}. Clustering and grouping of the samples were investigated with principal component analysis (PCA), while classification models based on FT-IR spectra were developed using the Soft Independent Modeling of Class Analogy (SIMCA) approach. The PCA plots showed clear separation between potstill brandies and blended brandies and some degree of clustering between the blended brands was observed. Classification tests resulted in a total correct classification rates between 81.25% and 100% for the various brandy brands included in this dataset. These preliminary results were very promising and highlight the potential of using FT-IR spectroscopy and multivariate classification techniques as a tool for rapid quality control in initial authentication of brandy brands.

5.1 INTRODUCTION

Brands are becoming an increasingly valuable asset for companies. To keep consumer preference, brandy producers aim to create a consistent product of conforming quality, vintage after vintage, and fiercely guard their intellectual property. Nevertheless, fraudsters use inferior, often potentially dangerous, raw materials to create products of inferior quality. In order to protect the integrity of the brand and assure quality to consumers, producers need fast, reliable and inexpensive techniques to aid the classification and authentication of brands.

South African alcoholic liquor law differentiates three different types of brandy: potstill brandy, blended brandy and vintage brandy. The former two types were used in this study. Potstill brandies are perceived as rich and layered in sensory evaluation and must have an ethanol content of at least 38 %v/v, while blended brandies are usually consumed with a mixer and must have an ethanol content of at least 43 %v/v (South African Liquor Products Act No. 60 of 1989). Vast differences in raw materials, production processes and price, make brand integrity and authenticity of utmost importance to producers and fuel the development of methods and techniques to protect and maintain it.

Traditional techniques used in brandy authentication studies, such as gas chromatography-mass spectrometry (Caldeira *et al.,* 2004), high-performance liquid chromatography (Canas *et al.,* 2003) and capillary electrophoresis (Panossian *et al*., 2001), focus on multi-component analysis of certain marker compounds and comparing the data with previously authentic samples' analyses (Sádecká *et al.,* 2009). These methods are time consuming, expensive and often require highly trained staff to interpret the results. Despite these drawbacks, these techniques are very sensitive and can detect trace amounts of marker compounds, frequently needed for absolute confirmation of authenticity of a product under question. However, given the high unit cost of analysis and time required per analysis, these techniques cannot be implemented for high sample through-put screening purposes.

Infrared spectroscopy and multivariate analysis have been used in various studies on alcoholic beverages, for authentication applications such as verification of brandy origin (Palma and Barroso, 2002), discriminating between tequilas (Lachenmeier *et al.,* 2005) and confirming brand identity of beer (Di Egidio *et al.,* 2011; Engel *et al.,* 2012). Near-infrared spectroscopy and chemometric methods were used (Pontes *et al.,* 2006) to classify and verify adulteration of distilled alcoholic beverages. Front face fluorescent spectroscopy and multivariate data analysis were used (Sádecká *et al.,* 2009) to classify brandies and wine distillates, offering a promising approach for authentication of brandies. Classification models using amino acid profiles of Chinese rice wines were developed (Shen *et al.,* 2011) with an overall correct classification rate of more than 90% for different ageing times and brands. Mid-infrared spectroscopy and chemometric techniques proved successful in the detection and quantification of adulterated Turkish distilled beverage, raki (Yucesoy and Ozen, 2013). Whisky samples were successfully classified into brands with chemometric profiles generated with gas chromatography-mass spectrometry results (Serino and Hom, 2013). Although infrared studies have become popular for authentication studies due to its simplicity and wide applicability, there has been little research carried out on brandy authentication in South Africa with this technique.

The aim of this study was to classify brandy products according to their brands and also directly assessing and assuring the quality of the branded product. Samples were sourced from the producing cellars and scanned with a FT-IR spectrometer. The spectra were analysed with PCA and SIMCA to classify the specific brands of brandies. This technique may allow authenticity testing simultaneously with quantification of specific sample parameters, such as ethanol, density and obscuration, when multivariate calibration algorithms are available on the instrument.

5.2 MATERIALS AND METHODS

5.2.1 Samples

Samples were collected at Distell (Distell Group (Pty) Ltd) routine laboratories. A total of 259 samples were collected during 2005, 2006 and 2009. The set included 38 potstill brandy samples and 221 blended brandy samples. The sample set included 12 Chateau VO, 38 Fish Eagle, 38 Klipdrift Export, 19 Klipdrift Premium, 12 Mellow Wood, 35 Olof Bergh, 16 Oude Meester, 61 Richelieu and 28 Viceroy samples. The spread of samples per brand is not ideal, but, since this is an industrial situation, the samples submitted to the laboratory were used.

5.2.2 Spectra

The spectral data was collected from 926 to 5012 to cm^{-1} and 1056 data points were collected per spectrum with a WineScan FT120 instrument (Foss Analytical, Denmark) as described in Chapter 4. Duplicate scanned spectra were first exported to the Unscrambler Software (version 9.2, Camo ASA, Trondheim, Norway) and averaged at each wavenumber to give a single spectrum per sample. The data was also exported from Unscrambler to the Simca-P^+ Software (version 13.0.2, Umetrics AB, Sweden). Spectral regions of water absorption were removed and wavenumbers trimmed to exclude noise in further analysis. The wavenumbers used in this study were 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}, giving 642 data points per spectrum.

5.2.3 Multivariate data analysis

Simca Software specific notifications were used throughout this report. Principal component analysis (PCA) was performed to investigate interpretable structure in the data (Simca User Guide, 2008). This technique reduces the entire spectral dataset to a small number of latent variables or principal components (PCs). Distance to model in X space (DModX) plots and Hotelling T^2 statistics (95% confidence level) were used to identify outliers (Simca User Guide, 2008).

Classification using the SIMCA approach was used to construct classification models for each of the brandy brands. In SIMCA, a PCA model is independently modeled for each class. A useful tool for the interpretation of SIMCA results is the Cooman's plot, which visualizes the discrimination between two classes (Berrueta *et al.,* 2007; Simca User Guide, 2008). SIMCA has been used in the classification and verification of adulteration in whisky, brandy, rum and vodka study with 100% success rate (Pontes *et al.,* 2006). A study on beer characterization

proved SIMCA as the best performing modeling technique in a comparison study with other modeling techniques (Di Egidio *et al.,* 2011). In this study nine models were defined. Models were evaluated by looking at the R^2 as a measure of goodness of fit, Q^2 as a measure of goodness of prediction. The optimal number of components for each model was selected, by evaluating the accumulated Q^2 value, so as to maximise the prediction capabilities (Simca User Guide, 2008). Results are scale dependent (Simca User Guide, 2008; Craig *et al.,* 2006), therefore the effects of different scaling options were tested on the classification models' accuracies. Mean-centering and no scaling, unit variance (UV) and Pareto scaling was evaluated for best fit to the data set for the classification models.

5.3 RESULTS AND DISCUSSION

5.3.1 Explorative data analysis

PCA is an unsupervised pattern recognition method of analysis. Therefore the brand of the sample was not taken into account, allowing for an overview of the data set. Hotelling T^2 (95% confidence level is indicated by the ellipse on the PCA plots. The samples inside the ellipse were considered compliant at a 95% confidence level. The wavenumbers used were 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}.

In Figure 5.1, the score plot is shown of PC1 and PC2 that respectively explained 93.4% and 4.7% of X data variance. The samples were coloured according to brand in Figure 5.1 (a) and brandy type in Figure 5.1 (b). Substantial overlapping between brandy brands was noted in Figure 5.1 (a), however a clear separation was seen between potstill brandies and blended brandies in Figure 5.1 (b). This clear differentiation between potstill and blended brandy types was expected due to the vast difference in production processes of the two types of brandy, as well as chemical composition and organoleptic evaluations.

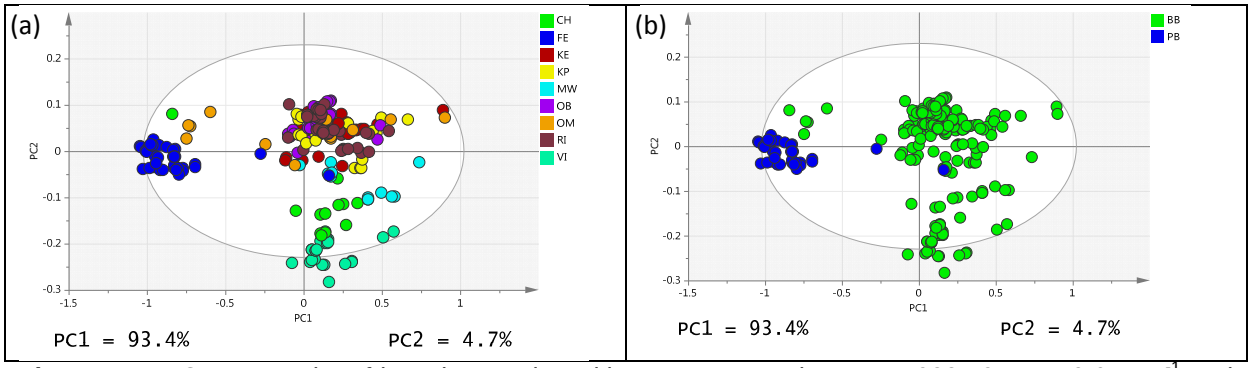


Figure 5.1: PCA score plot of brandy samples with mean-centered spectra, 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}. PC1 and PC2, explained 93.4% and 4.7% of the total variance respectively. The samples were coloured according to (a) brandy brand - Chateau VO (CH), Fish Eagle (FE), Klipdrift Export (KE), Klipdrift Premium (KP), Mellow Wood (MW), Olof Bergh (OB), Oude Meester (OM), Richelieu (RI) and Viceroy (VI) and (b) brandy type - Potstill brandy (PB) and Blended brandy (BB). Ellipse on the score plots indicate Hotelling T^2 range at 95% confidence.

In an attempt to investigate the blended brandies further, PCA was performed to explore the heterogeneity of the data set. The variable included in this study and that will be explored further, was year of spectral acquisition or scanning. Although various other variables such as production site, analyst, scanning instrument and stage of production at which the samples were drawn, also pertain to the samples, these variables were not investigated further, due to a lack of information.

PCA was performed on the samples scanned in one year separately, to examine qualitative differences within the year of scanning.

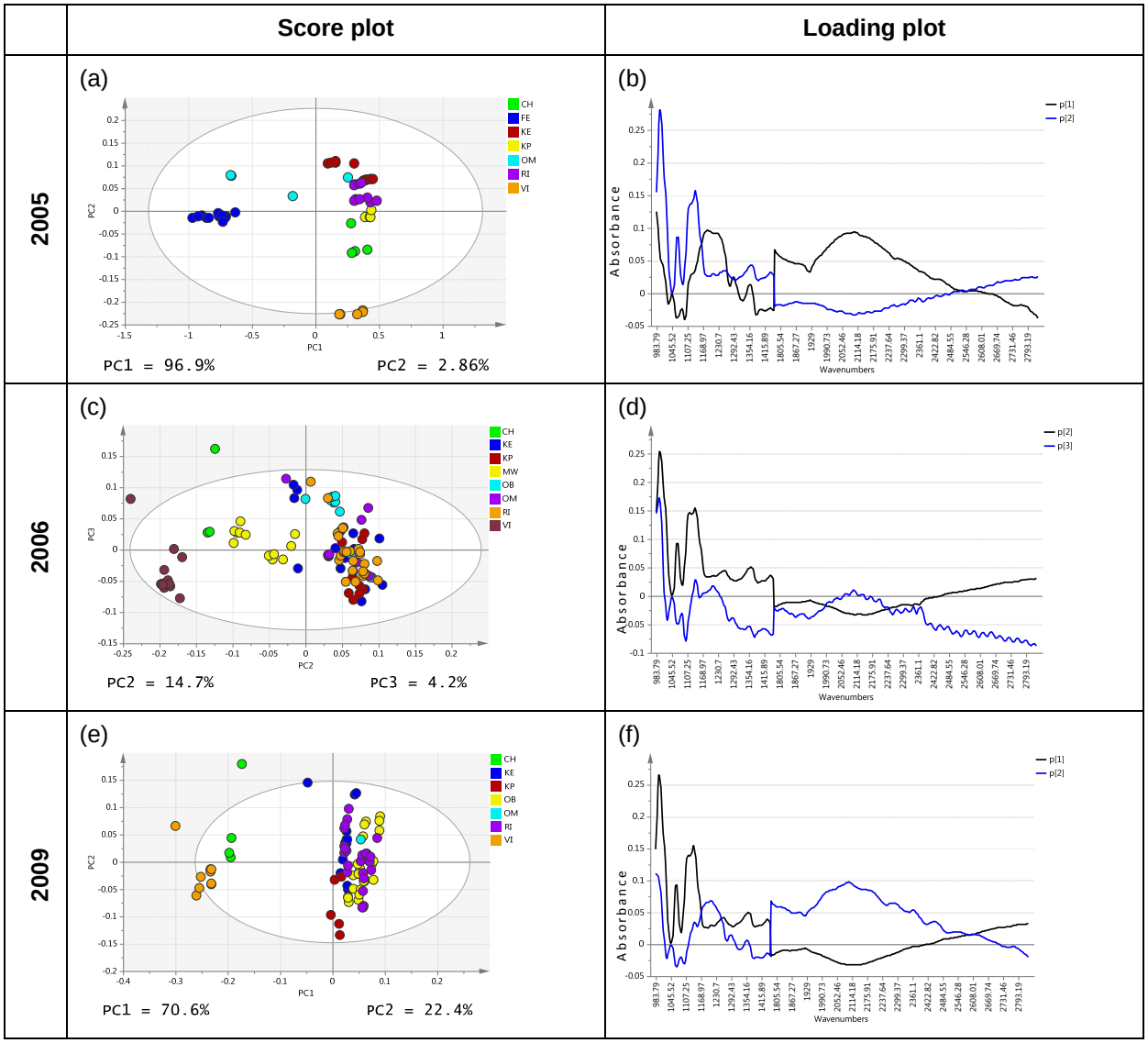


Figure 5.2: PCA score plot of brandy samples scanned in 2005 (a), 2006 (c) and 2009 (e). Spectra, 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1} were mean-centered. The samples were coloured according to brandy brand - Chateau VO (CH), Fish Eagle (FE), Klipdrift Export (KE), Klipdrift Premium (KP), Mellow Wood (MW), Olof Bergh (OB), Oude Meester (OM), Richelieu (RI) and Viceroy (VI). The loadings plots for each PCA are given next to the score plots. Ellipse on the score plots indicate Hotelling T^2 range at 95% confidence.

The score plot in Figure 5.2 (a) show that PC1 and PC2 respectively explains 96.9% and 2.86% of the X data variance, in the 2005 subset. Defined groupings related to the brand of the products, could be seen. In Figure 5.2 (b) the loading plot indicated the most important regions of wavenumbers responsible for the spread in Figure 5.2 (a). The peaks observed around 999.22 and 1138.11 cm^{-1} are areas associated with sugars and ethanol absorbances and contributed significantly to the spread seen in Figure 5.2 (a). These regions were identified in other studies as significant for separation in organic and non-organic wine (Cozzolino *et al.,* 2009), varietal separation of wines (Bevin *et al.,* 2008) and verification of adulteration in alcoholic beverages (Da Costa *et al*. 2004). In the case of the 2006 samples, principal components higher than PC1 were investigated, to try and interpret the clustering of the brands, since the strong contribution of ethanol on PC1, masked the contribution of other variables. Since the 2006 subset only contained blended brandy samples with an average ethanol concentration of 43 %v/v, projection on PC1 did not show clustering between the brands. PC2 and PC3 explained 14.7% and 4.2% of the data variance respectively. Loadings for wavenumbers between 983 and 1149 cm^{-1} were high on PC2 and PC3 and this region forms part of the fingerprint region (Nieuwoudt *et al.,* 2004; Alonso-Simón *et al.,* 2004), which is characterised by the absorption of many covalent bonds of organic chemical molecules of alcoholic beverages. The score plot, PC1 and PC2, was shown for the 2009 subset in Figure 5.2 (e), explaining 70.6% and 22.4% of the data variance respectively, and yielded similar results with regards to the wavenumbers responsible for the clustering as the 2006 subset Figure 5.2 (f).

Further investigation into the influence of year of scanning of the samples, was done with PCA. Within a brand the different year of scanning clearly had an influence on the positioning of the samples, as shown in the score plot (Figure 5.3), where Klipdrift Premium brandy samples were used as illustration. The 2005 and 2009 samples separated on PC1, while the 2006 samples separated on PC2. Therefore PC1 and PC2 loading plots were investigated. The wavenumbers that contributed to this separation were identified in the loading plots. It was not possible to interpret the sources of variation that contributed to the separation of year of scanning. This finding must be investigated further to explore the differences in the year of scanning and to standardise the product fingerprint over consecutive years.

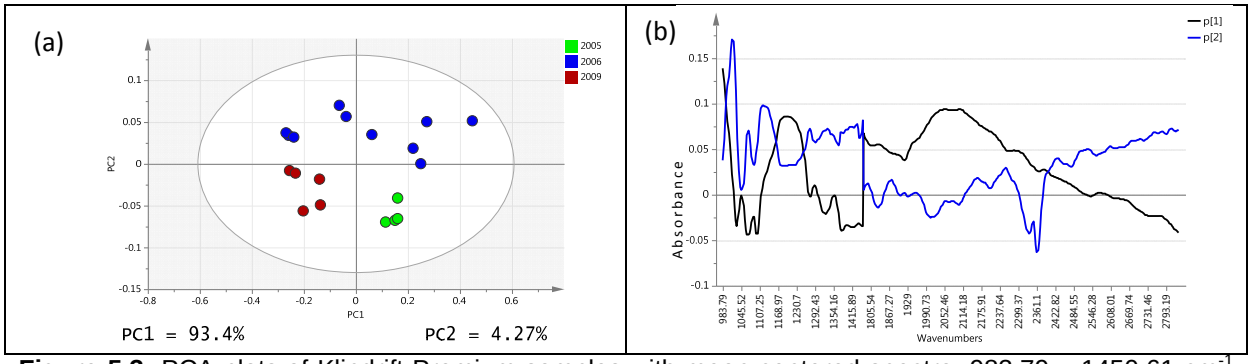


Figure 5.3: PCA plots of Klipdrift Premium samples with mean-centered spectra, 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}. The score plot of Klipdrift Premium samples scanned in 2005, 2006 and 2009 is given in (c) loadings plot for PC1 and PC2 (d). Ellipse on the score plot indicates Hotelling T^2 range at 95% confidence.

Although the brandies are created to specifications to keep the brands organoleptically unique and consistent between vintages, the products remain of natural origin, and vintage changes will ultimately have an effect on the end product. Palma and Barroso found that brandies of Jerez could be classified according to degree of ageing with an 83% chance of correct classification (Palma and Barroso, 2002). The effect of instrument drift could also play a role in the differences seen between years of scanning. Although the instruments on which the spectra were generated, are maintained under strict quality control procedures in the implicated laboratories, care should be taken and data interpretation should be corrected and compensated for using appropriate statistical and multivariate methods.

Groupings could be seen in PCA, but the separation of brandy brands was not clear and overlapping was observed between the brands. Therefore, further multivariate data analysis methods were applied to attempt classification of the different brandy brands.

5.3.2 Classification of brandy samples

To classify the observations according to brandy brand, different scaling options of the spectra were used. Three scaling treatments, namely no scaling and mean-centered, UV and Pareto scaling, were evaluated. A comparison of the classification results of the different scaling options is shown in Table 5.1, and it is clear that mean-centering provided the best option for this data set with an average of 91.89% correct classification over the nine brandy brands. The no scaling and no centering classification model yielded the lowest correct classification statistics of the brands (results not shown). Spectra without scaling are often dominated by peaks arising from a few strong spectral intensities which can obscure smaller but important features of the spectra (Noda, 2008). UV scaling on the other hand tends to amplify the effect of noise contribution, while Pareto scaling retains a small portion of the magnitude information of the spectral intensity variations (Noda, 2008). These results show the capability of the SIMCA model to successfully classify these brandies into their respective brands. These results are based on cross validation. The number of components was selected based on the highest explained variance with the lowest number of components to combat overfitting.

SIMCA allows for modeling per class. Each model per brand with mean-centering was evaluated separately in an attempt to improve the prediction ability by removing outliers. The model statistics, score plots (a) and DModX plots (b) are given in Table 5.2. The R^2 for each model was evaluated for goodness of fit, while the Q^2 was evaluated for the goodness of prediction and to prevent overfitting of the model, choosing the correct components per model (Simca User Guide, 2008). In the case of classification the dependent variable is a dummy variable.

Table 5.1: Comparison of different scaling methods on the classification of brandy brands using FT-IR spectra, 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}.

Brand	Number of samples	Mean-centered and no scaling	UV[a]	Pareto
Chateau VO	12	100%	100%	100%
Klipdrift Export	38	81.58%	76.32%	47.37%
Klipdrift Premium	19	100%	94.74%	84.21%
Oude Meester	16	81.25%	87.5%	93.75%
Richelieu	61	95.08%	73.77%	57.38%
Viceroy	28	92.86%	100%	96.43%
Mellow Wood	12	83.33%	100%	100%
Olof Bergh	35	94.29%	91.43%	91.43%
Fish Eagle	38	94.74%	100%	100%
Total	**259**	**91.89%**	**88.03%**	**79.15%**

[a]UV – Unit variance

The Hotelling T^2 ellipse in the score plots (a) and DModX plots (b) indicated a few possible outliers. These differences within a brand could be due to year of scanning, stage of production the sample was submitted to the laboratory, different machine the sample was scanned with or other variables included in this data set. Although some samples were flagged per brand, since this was an industrial feasibility study, all the samples submitted to the laboratories were used in the modeling and indicated the potential of the technique. In future, care should be taken to apply sample selection. Naes *et al.* (2004) emphasises the importance of sample selection for modeling.

Class separation for two models is shown in the Cooman's plot in Figure 5.4. The Cooman's plot does test set validation. In this case the rest of the data set and the Viceroy samples were tested against the Fish Eagle model. The x-axis shows the distance to the Viceroy class, while the y-axis shows the distance to Fish Eagle. The critical distances (at 95% confidence level) are shown with the DCrit (0.05) line in Figure 5.4. The figure shows successful classification and separation for these two models, and no misclassification of the prediction set into these two chosen models. This also confirms the clear separation between potstill brandies (Fish Eagle) and blended brandies (Viceroy) and rest of data set. This may provide a rapid method of brand determination.

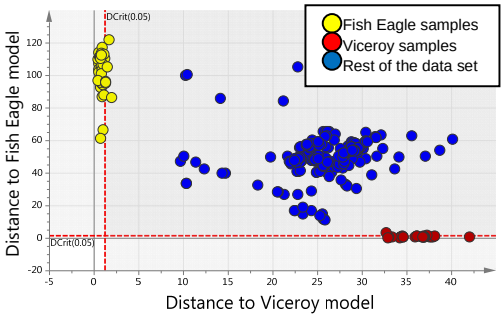
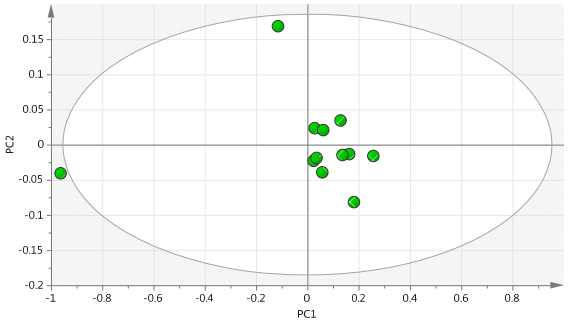
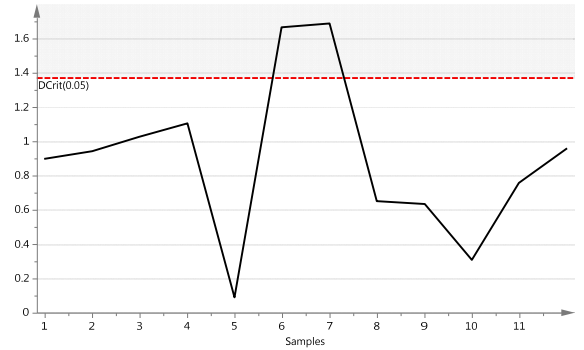
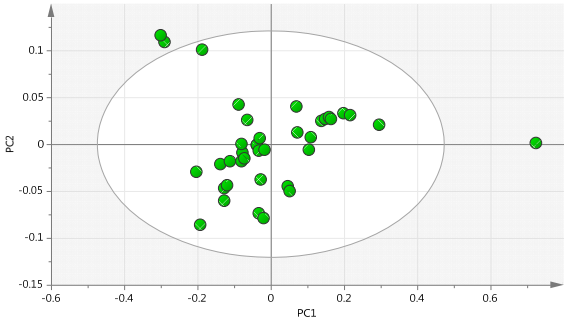
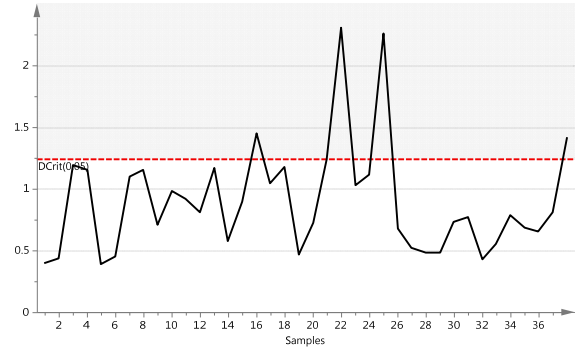
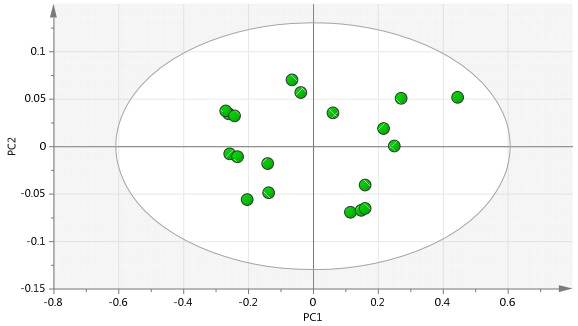
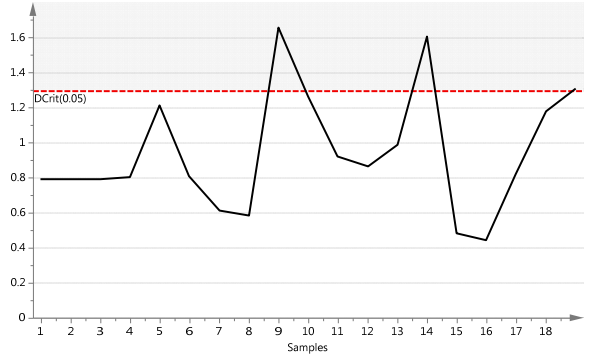
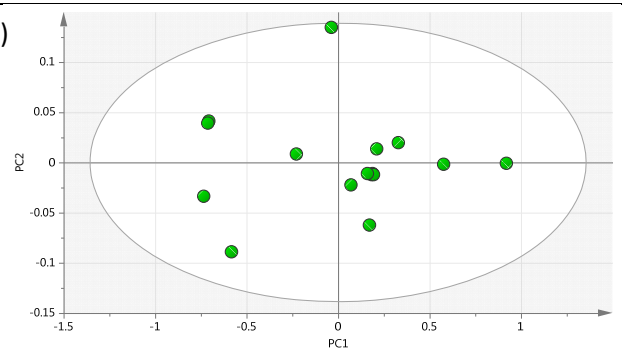
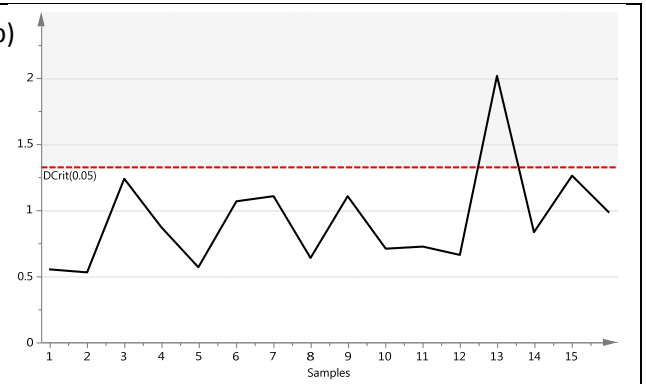
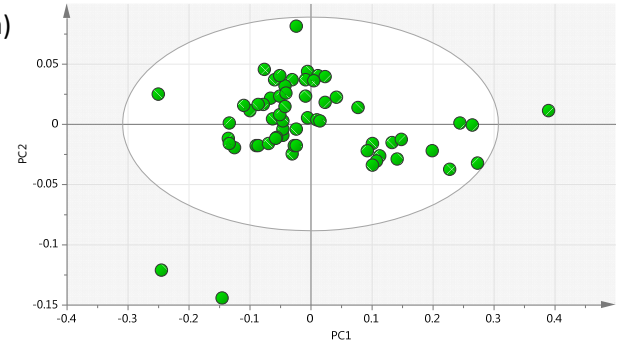
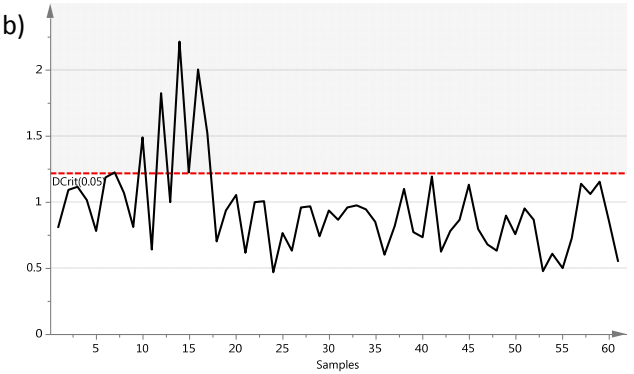
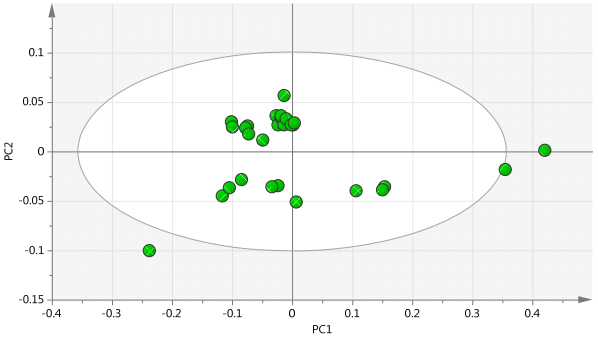
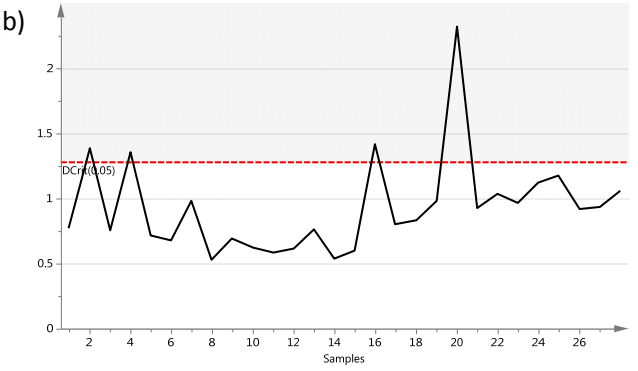
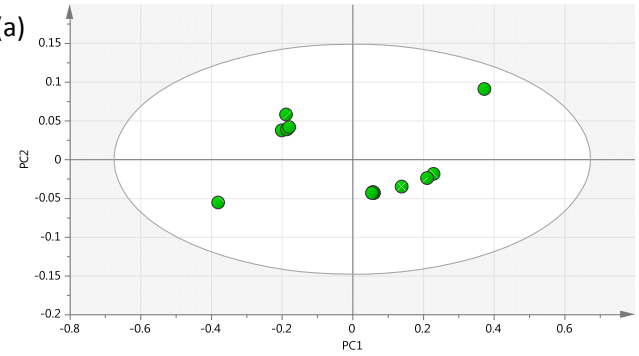
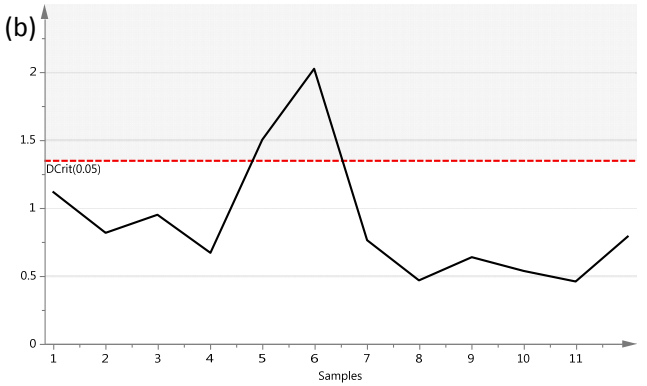
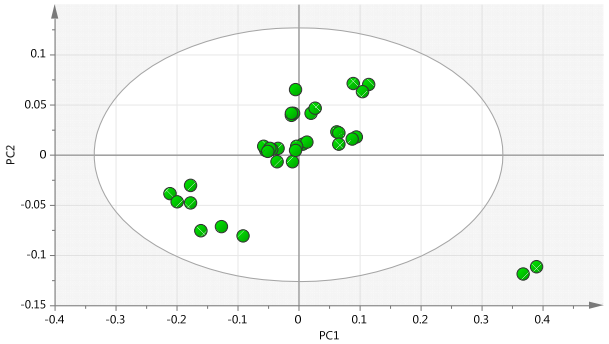
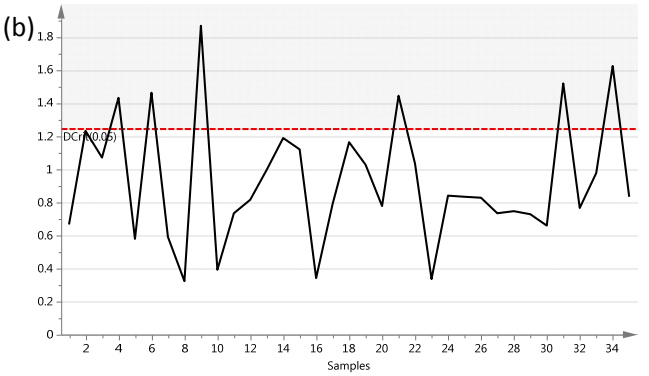
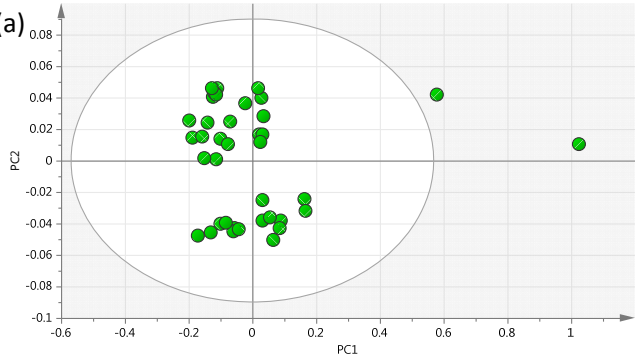
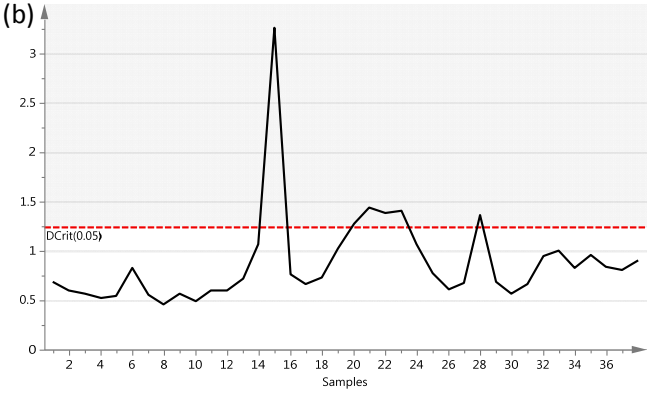


Figure 5.4: Cooman's plot for the classification of Viceroy (x-axis) and Fish Eagle (y-axis) samples using mean-centered spectra, 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}. DCrit at 95% confidence level is indicated with the red dashed lines.

Table 5.2: Summary of nine different brandy brands classification models based on SIMCA approach using mean-centered spectra, 983.79 – 1450.61 cm^{-1} and 1782.4 – 2835.63 cm^{-1}. Ellipse on the score plots indicate Hotelling T^2 range at 95% confidence. For each brand the model statistics, score plot (a) and DModX plot (b) are given.

CHATEAU Number of samples = 12 R^2 = 0.993 Q^2 = 0.949 Number of components = 3 PC1 = 93.3% PC2 = 3.53%	(a)  	(b) 
KLIPDRIFT EXPORT Number of samples = 38 R^2 = 0.98 Q^2 = 0.976 Number of components = 3 PC1 = 89.9% PC2 = 5.83%	(a)  	(b) 
KLIPDRIFT PREMIUM Number of samples = 19 R^2 = 0.977 Q^2 = 0.97 Number of components = 2 PC1 = 93.4%% PC2 = 4.27%	(a)  	(b) 

OUDE MEESTER Number of samples = 16 R^2 = 0.997 Q^2 = 0.993 Number of components = 3 PC1 = 98.1% PC2 = 1.03%	(a)  PC1 = 98.1% PC2 = 1.03%	(b) 
RICHELIEU Number of samples = 61 R^2 = 0.996 Q^2 = 0.994 Number of components = 6 PC1 = 87.7% PC2 = 7.26%	(a)  PC1 = 87.7% PC2 = 7.26%	(b) 
VICEROY Number of samples = 28 R^2 = 0.997 Q^2 = 0.987 Number of components = 5 PC1 = 88.5% PC2 = 7.03%	(a)  PC1 = 88.5% PC2 = 7.03%	(b) 

MELLOW WOOD Number of samples = 12 R^2 = 0.99 Q^2 = 0.979 Number of components = 2 PC1 = 94.4% PC2 = 4.57%	(a)  PC1 = 94.4% PC2 = 4.57%	(b) 
OLOF BERGH Number of samples = 35 R^2 = 0.995 Q^2 = 0.992 Number of components = 3 PC1 = 86.6% PC2 = 12.3%	(a)  PC1 = 86.6% PC2 = 12.3%	(b) 
FISH EAGLE Number of samples = 38 R^2 = 0.995 Q^2 = 0.993 Number of components = 3 PC1 = 96.6% PC2 = 2.41%	(a)  PC1 = 96.6% PC2 = 2.41%	(b) 

5.4 CONCLUSION

PCA performed on this data revealed a clear differentiation between potstill and blended brandies. These two types of brandies differ greatly in their production process, chemical parameters and organoleptically, which is supported by the PCA results. The PCA results also revealed the heterogeneity within the brands. Although specific brands are produced to specifications, inherent differences, that are not identified organoleptically or when looking at the chemical parameter results are clear in the spectra thereof, demonstrating the potential for quality control in brandy production. The loading plots of the PCAs indicate that the region between 983 and 1149 cm^{-1} had the most influence on the separation of the brands. The fingerprint region is important as the vibrational combination of the bands can be linked to single molecular bonds or functional groups that aid in the identification of compounds.
In this study the scaling treatments were evaluated and mean-centering proved to be the best fit for this data set. The mean-centered models showed a remarkable increase in classification ability. This method of scaling is most often used for spectroscopic data.

The performance of FT-IR spectroscopy and SIMCA method was evaluated for the classification of brandy according to their brand. Excellent classification results were obtained, however misclassification of brands was a risk. An average of 91.89% correctly classified brandy samples were obtained without removal of any outliers. These classification results have to be considered as preliminary, because of certain limitations in this study, such as the small number of observations per brandy brand and inherent variability in the data set, not modeled. The availability of a larger data set of authenticated brandy brands would further facilitate in the classification. Other supervised pattern recognition techniques may be explored in future studies for classification purposes.

Considering the encouraging results, the application of FT-IR in authentication studies, identifying falsified products by unscrupulous producers or to find markers to positively identify a brand seems to be a very promising approach. Development of a cost effective and mobile method of classification and ultimately authentication is possible using this study as a base.

5.5 REFERENCES

1. Alonso-Simón, A., Encina, A.E., García-Angulo, P., Álvarez, J.M. & Acebes, J.L. (2004). FTIR spectroscopy monitoring of cell wall modifications during the habituation of bean (*Phaseolus vulgaris* L.) callus cultures to dichlobenil. *Plant Science,* **167**, 1273 – 1281.
2. Bevin, C.J., Dambergs, R.G., Fergusson, A.J. & Cozzolino, D. (2008). Varietal discrimination of Australian wines by means of mid-infrared spectroscopy and multivariate analysis. *Analytica Chimica Acta* **621,** 19-23.
3. Berrueta, L.A., Alonso-Salces, R.M. & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A* **1158**, 196-214.
4. Caldeira, I., Pereira, R., Climaco, M. C., Belchior, A. P. & Bruno de Sousa, V. (2004). Improved method for extraction of aroma compounds in aged brandies and aqueous alcoholic wood extracts using ultrasound. *Analytica Chimica Acta* **513**, 125–134.
5. Canas, S., Belchior, A. P., Spranger, M. I. & Bruno-de-Sousa, R. (2003). High-performance liquid chromatography method for analysis of phenolic acids, phenolic aldehydes, and furanic derivatives in brandies. Development and validation. *Journal of Separation Sciences* **26**, 496–502.
6. Cozzolino, D., Holdstock, M., Dambergs, R.G., Cynkar, W.U. & Smith, P.A. (2009). Mid infrared spectroscopy and multivariate analysis: A tool to discriminate between organic and non-organic wines grown in Australia. *Food Chemistry* **116,** 761-765.
7. Craig, A., Cloarec, O., Holmes, E., Nicholson, J.K. & Lindon, J.C. (2006). Scaling and Normalization Effects in NMR Spectroscopic Metabonomic Data Sets. *Analytical Chemistry* **78**, 2262 – 2267.
8. Da Costa R.S., Santos S.R.B., Almeida L.F., Nascimento E.C.L., Pontes M.J.C., Lima R.A.C., Simões S.S. & Araújo M.C.U. (2004). A novel strategy to verification of adulteration in alcoholic beverages based on Schlieren effect measurements and chemometric techniques. *Microchemical Journal,* **78** (1) 27– 33.
9. Di Egidio, V., Oliveri, P., Woodcock, T. & Downey, G. (2011). Confirmation of brand identity in foods by near infrared transflectance spectroscopy using classification and class-modelling chemometric techniques — The example of a Belgian beer. *Food Research International* **44,** 544-549.
10. Engel, J., Blanchet, L., Buydens, L.M.C. & Downey, G. (2012). Confirmation of brand identity of a Trappist beer by mid-infrared spectroscopy coupled with multivariate data analysis. *Talanta* **99,** 426-432
11. Lachenmeier, D.W., Burri, P.A., Fauser, T., Frank, W. & Walch, S.G. (2005). Rapid determination of alcoholic strength of egg liqueur using steam distillation and oscillation-type densimetry with peristaltic pumping. *Analytica Chimica Acta* **537**, 377–384.
12. Næs, T., Isaksson, T., Fearn, T. & Davies, T. (2004). A User-Friendly Guide to Multivariate Calibration and Classification. NIR Publications, Chichester, United Kingdom.
13. Nieuwoudt, H.H., Prior, B.A., Pretorius, I.S., Manley, M. & Bauer, F.F. (2004). Principal Component Analysis Applied to Fourier Transform Infrared Spectroscopy for the Design of Calibration Sets for Glycerol Prediction Models in Wine and for the Detection and Classification of Outlier Samples. *Journal of Agricultural and Food Chemistry* **52**, 3726 - 3735

14. Noda, I. (2008). Scaling techniques to enhance two-dimensional correlation spectra. *Journal of Molecular Structure* **883-884**, 216-227.
15. Palma, M. & Barroso, C.G. (2002). Application of FT-IR spectroscopy to the characterisation and classification of wines, brandies and other distilled drinks. *Talanta* **58,** 265-271
16. Panossian, A., Mamikonyan, G., Torosyan, M., Gabrielyan, E. & Mkhitaryan, S. (2001). Analysis of aromatic aldehydes in brandy and wine by high-performance capillary electrophoresis. *Analytical Chemistry* **73**, 4379–4383.
17. Pontes, M.J.C, Santos, S.R.B., Araújo, M.C.U., Almeida, L.F., Lima, R.A.C., Gaião, E.N. & Souto, U.T.C.P. (2006). Classification of distilled alcoholic beverages and verification of adulteration by near infrared spectrometry. *Food Research International* **39,** 182-189.
18. Sádecká, J., Tóthová, J. & Májek, P. (2009). Classification of brandies and wine distillates using front face fluorescence spectroscopy. *Food Chemistry* **117**, 491-498.
19. Serino, T. & Hom, B. (2013). Chemometric Profiling of Whiskey using the 5977A GC/MSD. Agilent Technologies, Inc. Application note. Santa Clara, USA.
20. Shen, F., Ying, Y., Li, B., Zheng, Y. & Zhuge, Q. (2011). Multivariate classification of rice wines according to ageing time and brand based on amino acid profiles. *Food Chemistry* **129,** 565-569.
21. SIMCA. SIMCA-P+ 12 User Guide. UMETRICS, Umeå, Sweden; 2008.
22. South African Liquor Products Act No. 60 of 1989.
23. Yucesoy, D. & Ozen, B. (2013). Authentication of Turkish traditional aniseed flavoured distilled spirit, raki. *Food Chemistry* **141,** 1461-1465.

Chapter 6

General discussion and conclusions

CHAPTER 6 – General discussion and conclusions

Infrared spectroscopy coupled with multivariate data analysis techniques is an established method of analysis in international wine alcoholic beverage industries. Distell (Pty) Ltd, a major player in the spirit products industry in South Africa and internationally, needed to refine this technology to fulfil the requirements of a fast, accurate, cost-effective and reliable method of analysis and authentication of South African spirit products, to support quality control of their range of products. To address this need, three clearly defined objectives were identified for this study. These were (i) validation of the reference methods, (ii) development of PLS calibration models for the WineScan FT120 for ethanol, density, obscuration and colour in South African spirit products, and (iii) development of classification models to verify brands.

Regulatory compliance demands quality and reliability of results generated by chemical analysis (Taverniers *et al.,* 2004). For spirit products, the most important chemical/physical parameters analysed in routine analytical laboratories were identified and subsequently ethanol, density, obscuration and colour were the focus of this study. As a first step of the research, validation of the relevant reference methods was done. A distillation and oscillation-type density meter were evaluated for ethanol, density and obscuration analyses and a tintometer for colour analyses. The results obtained in the validation study, verified the use of these analytical techniques for the reference data collection, based on the accuracy and ruggedness achieved. The standard error of laboratory (SEL) throughout the validation study performed better than the SEL for a Distell laboratory (validated over three laboratories). For instance, the ethanol concentration error decreased from 0.200 %v/v for the Distell laboratory to 0.097 %v/v in this study. The differences in errors achieved could be attributed to the inclusion of various error adding variables in the Distell laboratory validation such as those contributed by different analysts, different equipment, different laboratory environments and measurements over time. The above finding that emerged from the results obtained in the method validation stage of the research, is common to error in measurement, but also pointed to the need to optimise infrared spectroscopy and multivariate data analysis as an alternative analytical technique, to minimise these errors, or at least have and additional measurement to validate the measurements obtained with chemical methods in future analysis.

The higher ethanol concentration and sugar in the spirit products than in wine, introduced new interferences to infrared spectroscopy and new calibration models had to be developed, adapting the existing ready-to-use calibration models, that were created with wine, available with the WineScan FT120 instrument, used to gather the spectra in this study. Fourier transform infrared (FT-IR) spectroscopy and partial least squares (PLS) regression were successfully used in calibration models for the analysis of ethanol, density, obscuration and colour of spirit products. The calibration models proved suitable for quantification and/or screening of these parameters for spirit products, thus further chemical analysis will be required for suspect or borderline samples. These were very robust models and incorporated future sources of variation. Gathering more samples per product type, can increase the variation within the dataset on which the calibration models are constructed and result in calibration models per product type suitable for accurate quantification. Although calibrations are transferable between instruments, some amount of slope and/or intercept adjustment is still required. Despite these adjustments, the error between instruments will still be significantly smaller and easily corrected than the errors between laboratories. This also highlights the importance of using accurate, precise and validated reference results upon which the regression models are built. Even with the remarkable progression in statistical packages, including erroneous reference results into the building of the regression model results in a less accurate model. This study has proved the

potential of this methodology application on spirit products in South Africa and offers considerable advantages above conventional reference analysis methods.

The generated spectra were used to evaluate FT-IR spectroscopy as a possible means of authentication of spirit products. Multivariate data analysis of the data acquired from scanning with an infrared spectrometer allowed samples to be separated successfully into product types with Principal Component Analysis (PCA). The PCA results confirmed that spectra contain information important for separation among products. They also suggest that FT-IR spectroscopy with multivariate data analysis holds the necessary information for successful classification of spirit product type. The interpretation of the loading plots indicated that these groupings could be as a result of differences in organic components, such as ethanol content, sugars as well as compounds containing aromatic rings, organic acids, aldehydes and esters. Clear groupings within a product type were observed, enabling preliminary indications of product identification and differentiation.

Nine brandy brands were used as the subset to evaluate FT-IR spectroscopy to correctly classify according to specific brand with Soft Independent Modeling Class Analogy (SIMCA) approach. The spectral region known as the fingerprint region 979.93 and 1450.61 cm^{-1} (Nieuwoudt *et al.,* 2004; Alonso-Simón *et al.,* 2004) was proven to contain useful information to identify brandy brands. The successful classification of the nine brandy brands in this study provides a more than reasonable base for proposing the use of this methodology for proving the authenticity of South African spirit brands.

A challenge exists in the use of infrared spectroscopy to authenticate spirit products and lies in the interpretation of the spectra obtained. Knowledge of the fundamentals of spectroscopy in addition to background information knowledge of the products and production processes were barriers in the authentication. Much time has been devoted in this study to the interpretation of PCA models to understand and substantiate the spread of the samples. The initial developmental stage of authentication methodology will involve much time to the identification of the wavenumbers of importance.

Future application of this technology could lead to the development of a fast and mobile device to determine authenticity of brands within the field, similar to the device described by Mackenzie and Aylott (2004), which used gas chromatography as methodology. Such a device would permit screening of samples in the field, consequently only the suspect samples could then be taken to the laboratory for a more in-depth analysis. Further, it is suggested to add organoleptic characteristics of the specific brands to the data set acquired, as to relate that to the spectra of the samples.

Unfortunately adulteration of spirit products can take various forms, such as addition of chemicals to compensate flavour or colour, substitution of labels with inferior products, production with inferior or undeclared raw materials, etc. Detecting the specific origin of adulteration will require an in-depth study. Evaluating the authenticity of a sample could simply be the comparison of spectra of the questionable sample to that of a guaranteed authentic sample, since the spectrum of a sample can aid as a chemical fingerprint, focussing on the region identified (983 and 1149 cm^{-1}) as important to separate products.

Overall, FT-IR spectroscopy coupled with multivariate data analysis could be a useful tool to guarantee the quality, preserve the image of the product and protect consumers against adulterated products and fraudsters, resulting in consumer loyalty and international establishment of South Africa as a front contender in the spirits industry.

REFERENCES:

1. Alonso-Simón, A., Encina, A.E., García-Angulo, P., Álvarez, J.M. & Acebes, J.L. (2004). FTIR spectroscopy monitoring of cell wall modifications during the habituation of bean (*Phaseolus vulgaris* L.) callus cultures to dichlobenil. *Plant Science* **167**, 1273 – 1281.
2. MacKenzie, W.M. & Aylott, R.I. (2004). Analytical Strategies to Confirm Scotch Whisky Authentication, Part II: Mobile brand authentication. *Analyst* **129**, 607 – 612.
3. Nieuwoudt, H.H., Prior, B.A., Pretorius, I.S., Manley, M. & Bauer, F.F. (2004). Principal Component Analysis Applied to Fourier Transform Infrared Spectroscopy for the Design of Calibration Sets for Glycerol Prediction Models in Wine and for the Detection and Classification of Outlier Samples. *Journal of Agricultural and Food Chemistry* **52**, 3726 – 3735.
4. Taverniers, I., Van Bockstaele, E. & De Loose, M. (2004). Trends in quality in the analytical laboratory. I. Traceability and measurement uncertainty of analytical results. *Trends in Analytical Chemistry* **23**, 480 – 490.